CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

THE EFFECTS OF TYPE OF INFERENCE ON CHILDREN'S ABILITY TO PERFORM INFERENTIAL REASONING TASKS

by

Phyllis Mary Joan Cassidy

B.A., Simon Fraser University, 1973

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS (EDUCATION)

in the Faculty

of

Education

SIMON FRASER UNIVERSITY

November, 1985

ISBN 0-315-30807-9

APPROVAL

Name: P. M. Joan Cassidy

Degree: Master of Arts (Education)

Title of Thesis: The Effects of Type of Inference on Children's Ability to Perform Inferential Reasoning Tasks

Examining Committee

Chairperson: Ronald W. Marx

P. H. Winne
Senior Supervisor

J. R. Kendall
Associate Professor

B. Wong
Associate Professor
Faculty of Education
Simon Fraser University
External Examiner

Date approved 30 October, 1985

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

THE EFFECTS OF TYPE OF INFERENCE ON CHILDREN'S ABILITY TO PERFORM INFERENTIAL REASONING TASKS

Author:

(signature)

Phyllis Mary Joan Cassidy

(name)

Oct. 30/85.

(date)

ABSTRACT

This research replicated and extended a study conducted by Lipson (1982). It investigated grade three students' comprehension of expository text in terms of their ability to recognize two types of information presented explicitly in text and to infer four types of information based on premises presented implicitly in text. The students were classified by their teacher as reading above, at or below grade level. They read aloud four short passages. After reading each passage, the students were administered four measures: free recall; a recognition task presented in two conditions, text absent and text present; and finally a premise identification task with text present.

Responses to the recognition tasks were analyzed by performing a 3(reading level) x 2(explicit vs. implicit information) x 2(text absent vs. text present) x 4(passage) multivariate analysis of variance. Reading level was the sole between subjects factor. Passage, type of information, and text absence or presence were the within-subjects factors. This analysis confirmed Lipson's findings that explicit information was more readily recognized than implicit information. It also revealed that performance improved only marginally when the text was available during question answering.

This finding suggests that memory deficiencies cannot be held accountable for poorer performance in inferential reasoning tasks. To determine whether type of inference affected performance a 3(reading level) x 4(inference type) x 4(passage) x 2(text absent vs. text present) multivariate analysis of variance was conducted. No main effects were found for either passages or presentation condition, however, reliable differences were found amongst levels of inferences. The data supported Lipson's findings that children recognized lower level attribute and goal inferences more readily than higher level event and causal inferences. Another question investigated was whether students' ability to identify premise information influenced their recognition of inferences. A 3(reading level) x 6(question type) x 4(passage) multivariate analysis of variance revealed that more premise information was correctly identified for attribute and goal inferences than for event and causal inferences. This study demonstrates that young readers' abilities to draw inferences are heterogeneous. Future research should seek to explain these findings. Teachers should be aware that some types of inferences are easier for students to make than are others, and that these differences bear on designing reading instruction and assessing readers' level of achievement.

DEDICATION

For my father, Mac Collinge,

and in loving memory of my mother,

Phyllis Collinge,

with appreciation, love, and the deepest respect.

ACKNOWLEDGEMENTS

I wish to begin my words of acknowledgement by thanking my Senior Supervisor, Dr. Phil Winne. The generous manner in which he made available his time and knowledge made the preparation of this thesis a pleasurable and rewarding experience. I also wish to thank the second member of my committee, Dr. Janet Kendall, for her comments, criticisms, and suggestions. Appreciation must also be extended to Dr. Jupian Leung who assisted so patiently and co-operatively with the analysis of data.

Don McDonald, Loretta Greeno, and the grade three children of Burquitlam Elementary School are to be thanked for their willing participation in this study, as are Brian Evans, Val Douglas and the grade three students of Anmore Elementary School.

To my family and to my family of friends, I express my very sincere appreciation. Your support, delivered in many different ways, made this task a far easier one.

The final words of gratitude I save for my daughter, Brighid, whose understanding, love and acceptance have been a source of pleasure and encouragement throughout. Thank you, Brighid. You have given more than I could ever have imagined possible from one so young.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

Reading comprehension is commonly defined as the process whereby meaning is acquired from written text. According to this view, reading is considered by many to be receptive and passive in nature. Such a view, however, has been criticized because it fails to acknowledge the complex processes that are associated with acquiring meaning. Another point of view is that espoused by Wittrock (1984). While it does not deem the other incorrect, it does identify its narrowness. Wittrock argues that reading comprehension is not a passive task, but rather the creative process of generating meaning for discourse composed by other people.

The appropriateness of conceptualizing reading comprehension as the process of generating meaning is apparent when one considers that text seldom, if ever, contains all the information required to understand what appears in print. Authors expect readers to "fill in" and "connect" information in certain predictable ways (Lipson, 1982). Readers, if they are to understand connected discourse, must construct their own relationships between various parts of text. Indeed, if authors were to fill in every piece of missing information, reading would be rendered a rather laborious endeavour. Observe, for example, the tedium of the following passage.

> James Bond lifted his left leg up and moved it forward quickly, putting his left foot in front of his right, touching the heel of his shoe to the ground and then

> the toe ... Grasping the ignition key firmly between his right thumb and forefinger, he moved his hand toward the ignition, inserting the key. He twisted the key to the start position, while pressing slightly on the accelerator with the toe of his right shoe.
> (Anderson, 1981, cited in Broatch, 1981, p. 4)

The more parsimonious version - "James Bond dashed out of the casino, jumped into his Maserati, and sped away" -- though "incomplete", is far more likely to retain the interest of the reader and, by omitting minute detail, better allow the reader to construct the author's intended message.

Text can be viewed as a set of directions to the readers' thought processes. From the cues provided, the reader must then construct mental representations consistent with the author's guidelines (Wittrock, 1984). According to Johnson-Laird (1983), without the ability to generate information not explicitly stated, written or spoken discourse would be beyond anyone's competence. Where skills in this area are impaired, so too is comprehension (Lehnert, 1981).

Inferencing is the label attached to the process of generating unstated information and relationships. It is defined by Frederiksen (1979) as those "processes that operate on <u>given</u> propositions to produce <u>new</u> ones" (p. 172).

From both a theoretical and empirical perspective, inferencing is considered a critical component of the comprehension process. Understandably, it has received considerable attention from reading researchers. A common deficiency of much of this research is that in

designing studies investigators often fail to adequately operationalize the complexity of the inferential process. Typically, inferences are dichotomously categorized, their classification being determined by whether the source of premise information is text-specific or text-external. It seems doubtful that a full appreciation of the complexity of inferential reasoning will be acquired from such studies. To limit classification of inferences to two such categories requires the assumption of homogeneity amongst all entries under each heading. As will be shown in Chapter Two, such an assumption is not justified.

In order to improve upon the shortcomings of such research, it would seem appropriate first to identify the different types of inferences required of readers and second, to test for varying degrees of difficulty each may present. One study that addressed these issues was Lipson's (1982). Her work is replicated and extended here. Lipson investigated children's comprehension of explicitly stated and implied information. Grade three students read short expository passages and responded to a set of paired recognition items constructed from information contained in the passages. The recognition items were exemplars of four types of inferences: attribute, goal, event and causal. The results suggested that certain inferences, namely event and causal, were more difficult for children to recognize than others.

Lipson's work is noteworthy on two other accounts. First, her

study examined inferencing behaviour in children rather than adults. As she points out, "much of the research on discourse processing lacks applicability for educational settings because almost all of it has used adult subjects. There is a need to examine the discourse processes employed by children who are not yet fluent readers" (Lipson, 1982, p. 245). Secondly, expository text was used rather than narrative. Of those studies that do investigate children's comprehension, most have used only narrative material.

There were two purposes in replicating Lipson's work. First, it was of interest to ascertain if similar findings would be obtained regarding the relative difficulty of different types of inferences. Secondly, by extending her methodology, questions left unanswered by the original study could be addressed. These questions will be discussed in Chapter Two.

CHAPTER TWO

Reading has been defined in many ways. At one extreme there is Thorndike's (1917) interpretation of reading as reasoning, at the other Flesch's (1955) conceptualization of reading as deciphering the printed code. These definitional differences are reflected in reading research where currently two types of research are pursued. One examines the perceptual processes involved in reading, the other the processes of comprehension (Marshall, 1981). The latter is currently one of the most active areas of research in psychology (Freedle, 1977, 1979; Meyer & Rice, 1979) and the topic that is addressed here.

For those who investigate comprehension there appears to be general agreement that reading can be defined as "a complex of mental processes that a reader uses when interacting meaningfully with printed discourse" (Marshall, 1981, p. 40). The processes to be discussed here are those involved in comprehending implicit information.

Before proceeding, the differences between implicit and explicit comprehension should be defined. Paris and Upton (1976) argue that the two are differentiated by the degree of effort or the level of processing required for comprehension to occur. Explicit comprehension, also referred to as literal or verbatim comprehension, involves extracting meaning from information explicitly stated in the text (Oakhill, 1984). Implicit comprehension, on the other hand,

requires constructing unstated relationships based on information explicitly present in the text (Paris & Upton, 1976).

The process of generating unstated propositions is called inferencing, and it is considered to be a vital and necessary component in comprehending written discourse (Flood, 1981; Paris, 1975; Reder, 1980). Inferences are "propositions not explicitly revealed in the text which readers conjure from their schemata or ferret from ideas stated in the text in order to fully understand and/or recall discourse" (Olson, 1985, p. 200). It is because text does not contain all that is necessary for it to be fully understood that the construction of inferences is so very important. Language, explains Spiro (1980), provides but a skeleton of meaning, and it is through the generation of inferences that the reader adds substance to the skeletal representations provided.

According to Johnson (1983), there are at least four ways in which inferences facilitate comprehension. One, they reduce lexical ambiguity. Two, they resolve pronominal and nominal reference. Three, they contribute to the cohesiveness of text by joining together propositions and sentences. Four, they fill in missing "chunks" of information by providing a larger framework from which to interpret the material presented. For these things to occur, readers are required to draw on both prior knowledge and textual content (Olson, 1985).

Textual content and prior knowledge represent the two end-points

of a continuum characterizing the informational source of inferences (Bartlett, 1932; Olson, 1985; Pearson & Johnson, 1978). It has been suggested that inferencing is initially fairly text-based in that the inferences drawn are mainly those triggered by discourse features, anaphoric reference for example. With the acquisition of greater skill the process becomes more, though not entirely, schema-based (Johnson, 1983). That is to say, prior knowledge structures are activated by the material being read. This facilitates comprehension by providing a context with which to interpret incoming information. Inferences of both sorts lie at the very heart of comprehension since it is generally believed that the more inferences a reader makes the more he comprehends (Johnston, 1983). The reason for this is that inferential processing stores in memory an integrated, semantically rich representation of the text (Paris & Lindauer, 1976).

Inferencing not only enhances comprehension, it also has been found to facilitate later recall (Oakhill, 1984). Paris and Upton (1976) explain this finding by suggesting that when "a person constructs implied relationships and integrates them with original information, the derived memory representation may permit a temporally ordered, logical, sequential unit that can be stored parsimoniously and accessed readily" (p. 667).

Inferencing, therefore, plays a critical role in comprehending written discourse and because of this has received much attention from reading researchers. The resulting literature is difficult to review,

however, because of the inconsistent manner in which inferences are labelled across studies. Without attempting to be exhaustive, a few examples will demonstrate the difficulties encountered. Before beginning, however, it is important to note that the distinction between implicit and explicit inferences about to be elaborated should not be confused with the more general distinction between implicit and explicit information. The latter denotes the difference between unstated and stated information in text. The former refers to two different kinds of inferences.

To begin, consider Johnson-Laird's (1983) definition of implicit inferences. These, he explains, are formed rapidly and without effort. They "underlie the more mundane process of intuitive judgment and comprehension" (Johnson-Laird, 1983, p. 127). Oakhill (1984), however, provides another interpretation. She states that implicit inferences "make sense of a situation in terms of relevant general knowledge, i.e. all the premise information is not presented" (p. 31). Such explanations provide insufficient detail to determine if the investigators are referring to the same thing. The confusion increases when these definitions are juxtaposed with those describing explicit inferences. From Johnson-Laird's point of view, explicit inferences are those which result when conscious effort is applied to connect information in text with that available from prior knowledge. The necessity to draw on prior knowledge, however, is what Oakhill deems an important feature of implicit inferences. She states that

explicit inferences are derived directly from the information presented in text. She makes no mention of the need to activate prior knowledge in generating explicit inferences. Both Johnson-Laird and Oakhill agree, however, that explicit inferences are the product of deliberate deduction.

To confuse the issue further, Paris and Upton (1976) introduce the terms "lexical" and "contextual" inferences. The former refer to those that are "constrained primarily by single lexical items" (p. 661), the latter to generating relations either within or between sentences. Here, no mention is made of the presence or absence of all premise information, or the need to call upon prior knowledge.

Further confusion over what inference making is can be illustrated from Frederiksen's (1979) work on text-based inferences which he posits occur "primarily in response to discourse features" (p. 165). He compares text-based inferences to schema-based inferences, the latter being those which are generated when a reader activates prior knowledge structures. Text-based inferences are similar to Johnson-Laird's implicit inferences in that they are formed rapidly and without effort. They differ from Oakhill's implicit inferences since she proposes that implicit inferences are generated from prior knowledge. Frederiksen would argue that all information required to construct text-based inferences can be found in text.

Some of the confusion over distinctions among types of inferences can be attributed to the difficulty of precisely identifying the

source of premise information. This appears to be one important criterion in determining the label applied to inferences. The issue of determining the source of information is difficult to resolve. Even for those inferences that are generated rapidly and without reflection and that are based on text that contains all necessary premise information, it would be hard to imagine that some prior knowledge is not activated in their construction. Thus, some inferences, such as anaphoric references, can be easily distinguished as text-based; others can be easily distinguished as schema-based, such as those constructed during the reading of the James Bond passage in Chapter One. Where other inferences may be positioned along the text-based/schema-based continuum is very much a product of the previous experience and world knowledge a reader brings to the task of reading.

Another reason for the labelling confusion, and one more encompassing than that noted above, may be that both the theory and data needed to provide a well-specified definition of inferences is lacking, as is that which would delineate how inferences may differ from one another (Paris & Upton, 1976; Warren, Nicholas, & Trabasso, 1979). Research efforts, however, have been applied to overcome this deficiency, with the result that various systems for classifying inferences have been proposed (Clark, 1977; Flood, 1978; Frederiksen, 1979; Warren, Nicholas, & Trabasso, 1979). These systems will be presented now in abbreviated form with a brief comment regarding their

suitability for research investigating children's comprehension of written expository discourse.

Clark's system was developed to classify inferences in spoken discourse. The mutually exclusive categories of inferences established were based largely on rules of conversation and are not related to the structure of text. As the relationship between oral and written discourse remains unclear, Clark's system is considered inappropriate for research on prose comprehension (Warren, Nicholas, & Trabasso, 1979).

The inference taxonomy developed by Warren, Nicholas and Trabasso (1979) was designed to describe inferences readers make while reading narrative text. Within this system inferences fall within one of three categories. One, logical inferences include the motivations, causes and conditions that enable an event to occur. They are the inferences drawn in response to "why?" and "how?" questions. Two, informational inferences answer "who?", "what?", "when?" and "where?" questions by providing information about specific people, instruments, objects, times, places, and contexts of events. Three, value inferences are evaluative in nature and reflect the reader's judgements on the story. The taxonomy developed by Warren et al. is valuable because it allows researchers to define inferences with greater precision than the dichotomous groupings described earlier in this chapter. However, because it was developed to reflect inferences generated in response to narrative text and because it is not known

whether both narrative and expository text prompt the same kinds of inferences (Olson, 1985), Warren et al.'s taxonomy is considered inappropriate for a study examining comprehension of expository passages.

In the taxonomy of inferences developed by Flood (1978) six categories of inferences are proposed. There are those generating macro- and micro-structures whereby larger or smaller units are created to accommodate information in text. For example, in recall the word "flower" may be substituted for the content word "daffodil". Other types of inferences are those generating cause, dimension, case frames, and attributes and finally those which establish appropriate references for ambiguous elements of text. This taxonomy was based on a study which analyzed a single text and reflects the type of processing in which competent adult readers would engage. As Lipson (1982) indicates, however, fluent and less fluent readers may process information differently. Therefore, inferences generated by college freshmen may not necessarily reflect the type of inferences made while children read.

Frederiksen's (1979) work overcomes the limitations of the taxonomies of inferences described above. First, the categories of inferences reflect the underlying structure of text. Second, the taxonomy was designed for use with both expository and narrative material. Third, in formulating the taxonomy, specific attention was directed to the types of inferences found in children's recall

protocols. Because Frederiksen's work is adopted in this research, it will be described in greater detail.

Frederiksen's Research on Reading Comprehension in Young Children

Paris and Upton (1976) stated that classifying inferences is a difficult undertaking, partly because neither inferences nor the semantic differentiations between them have been well specified. They argue that given the flexibility of semantic expression found in language, it would be difficult to establish rigidly defined categories. However, as is evidenced from the discussion regarding taxonomies of inferences, several theoreticians have attempted to do just that. The taxonomy to be discussed now was developed by Frederiksen (1979). His theoretical work has made two contributions to the field of prose comprehension: one, a system for representing the propositional structures underlying the comprehension of discourse; and two, a taxonomy of text-based inferences. Although the purpose of the present study deals not with text representation but inferencing, a brief description of Frederiksen's system for representing knowledge will be presented as it facilitates understanding some aspects of the taxonomy of inferences.

Knowledge Representation

In order to examine inferential processing, Frederiksen considered it necessary to develop a system for representing the

information upon which these processes operated. The system was designed to model not only the structure of information in text, but also that in long-term memory. He argues that a "text's structure is a reflection of the knowledge structure of the speaker or writer who produced the text ... if a semantic distinction is manifest in language, it must also represent an aspect of human memory structure" (Frederiksen, 1977, p. 57).

Frederiksen hypothesizes that networks provide the organizational structure for semantic information in human memory. A network is composed of nodes (ideas or concepts) and links (the connections which specify the relationship between two nodes).

Frederiksen (1975) intended to develop a system which would have "general applicability as a model of memory structures from which speakers (or writers) generate English texts" (p. 378). With this goal in mind, he deemed it necessary and therefore attempted to account for (a) every concept which is "lexicalized" in English, and (b) every relationship that could exist between any two concepts. His list will not be elaborated upon here as it is not necessary for the purposes of this study. It is important to remember, however, the two basic elements in his system: concepts and relations. It is from these that all other units within the semantic networks are constructed.

Frederiksen's system is composed of six units which are listed and defined in Table 1. They are ranked in order of increasing

Table 1

Ranked Units in Semantic Structure

Concept	Objects, actions, attributes of objects or actions, numbers, degrees, locations, and times.
Relational Triple	Two concepts joined by a labelled relaitonship. The types of relationships found at this level are: case relations (which specify the way in which nouns are related to verbs); stative relations (which describe an object); determiner and quantifier relations; manner relations; and degree, locative and temporal relations.
Event Frame	An action plus a network of case relations. Unlike objects, actions cannot stand alone. They always involve causes and effects. Thus there is a needed frame into which can be placed the action and its causes and effects. Event frames, however, do not represent a complete event. Time, location, and other characteristics of an event are not included here. This information is contained in a proposition. An event frame, therefore, is a component of a proposition.
Proposition	A state or an event. A state is an object plus a set of stative relationships that identify that object. An event is an action plus a set of relations. These relations can be of two kinds: (a) case relations and (b) identifying relations.
Relative Systems	Two relative or metric propositions joined by an algebraic relation or function. Algebraic relations connect propositions which contain identical transitive (ordered) or non-transitive (non-ordered) relations. They specify relative time, location, or comparative information involving attributes of objects or actions (Frederiksen, 1979).
Dependency Systems	In this system one proposition is related to another in a manner which renders one dependent on the other. A proposition can be dependent on another either logically, conditionally, or causally.

complexity, commencing with the smallest unit, the concept, the only one which does not subsume another unit in its makeup. It proceeds to the largest and most inclusive unit, the dependency system (Frederiksen, 1977).

Frederiksen's system of propositional representation is considered to be particularly useful when the purpose of research is to examine the inferences made by children while reading. Because the system was designed to study inferences, Frederiksen made finer distinctions among relationships than any other system of text representation (Meyer & Rice, 1984). The system, however, is very intricate and goes beyond the depth of analysis required for the present study. The summary provided here is intended to provide the necessary background to better understand the ways in which inference types can be differentiated from one another.

The concepts underlying Frederiksen's system of text representation having been described, the discussion will proceed to the second theoretical foundation of his theory of discourse comprehension -- the taxonomy of text-based inferences.

Taxonomy of Text-based Inferences

Frederiksen's approach to text analysis provided the foundation from which evolved an exhaustive listing of the types of inferences contained in children's story recalls. The units of information outlined in the previous section are the units upon which the

inferencing processes operate. The inferencing processes fall into eight categories. One of two things occurs when an inference is made: a particular process operates on different semantic units; and/or a single semantic unit is acted upon by different cognitive processes. A process operating on units creates an inference.

The taxonomy, considered by Frederiksen to be exhaustive, was developed by analyzing two sources of information: one, the types of inferences he considered possible, given the nature of propositional structures; and two, the types of propositions found in young children's story recalls. Table 2, borrowed from Frederiksen (1979, p. 173), lists the 26 major types of text-based inferences and specifies the units of information acted upon by each class of operations.

Empirical Investigations of Inferencing

Unfortunately, much of the applied research on inferencing does not draw on the various theoretical efforts just described. Many studies, for instance, limit themselves to investigating differences between the comprehension of explicit and implicit information. Others, while acknowledging the existence of different types of inferences, differentiate them dichotomously, often categorizing inferences as either implicit or explicit, text-based or schema-based, lexicalized or contextualized. Such studies can provide valuable information and their importance should not be denied. For diagnostic

Table 2

Major Classes of Text-Based Inference

Classes of Operations	Units	Inference Types
I. Lexical operations	Lexical concepts propositions	1. Lexical expansion 2. Lexicalization
II. Identification operations	Objects, actions states, events	3. Attribute inference 4. Category inference 5. Time inference 6. Locative inference 7. Hasp inference 8. Degree inference
III. Frame operations	Event frames	9. Act inference 10. Case inference 11. Instrumental inference 12. Result inference 13. Source inference 14. Goal inference 15. Theme inference 16. Frame transformation 17. Disembedding operations 18. Embedding operations
IV. Event generation	Event propositions	19. Event inference
V. Macrostructure operations	Propositions	20. Superordinate inference 21. Subordinate inference
VI. Algebraic operations		22. Algebraic inference A. Metric B. Nonmetric 23. Causal inference
VII. Dependency operations	Dependency systems	24. Conditional inference A. Enablement inference B. Presupposition C. Antecedent inference

Table 2 (Continued)

Classes of Operations	Units	Inference Types
		25. Logical inference A. Deductive inference B. Conditional perfection
VIII. Truth-value operations	Propositions	26. Truth-value operations A. Qualification B. Negation

Note: From Discourse comprehension and early reading by C. H. Frederiksen. In L. B. Resnick & P. A. Weaver (Eds.), Theory and practice of early reading: Vol. 1 (p. 173). Hillsdale, N.J.: Erlbaum. Copyright. 1979.

and prescriptive purposes in teaching, for instance, it can be advantageous to know what kind of information or what type of inference is processed more readily or causes greater difficulty.

Having acknowledged the worth of such studies, it is equally important to note their shortcoming, namely their failure to account for the heterogeneity of inferences grouped under a generalized heading. As is evident from the discussion of Frederiksen's taxonomy, an assumption of homogeneity amongst types of inferences identified only by a dichotomy probable is not justified. Not only can the unit of information processed in generating an inference vary within one category, but the cognitive processes that act upon those units also may differ. By way of example, consider the second major class of inference in Frederiksen's taxonomy, identification operations. These operations can act upon both objects (a flower pot, for instance) and actions (going to the movies). Depending on the nature of the question and the processing triggered, any one of six types of inferences can be produced. For instance, the attribute "new" may be attributed to the flower pot, resulting in an attribute inference, or the object "flower pot" may be classified as a vase, in which case a category inference has been generated. Both inferences would be classified as text-based, and in both cases the object acted upon is identical. But the cognitive processes operating on that object differ and give rise to a different type of inference. If we are to better understand the inferential processes of children, it may be

necessary to take such factors into account when designing research.

A study conducted by Lipson (1982) was a positive departure from the type of research described above because it investigated children's performance on four different types of inferences. Lipson was interested in examining the ability of grade three children to acquire new information from expository text. To assess comprehension of the material read she tested performance on four types of inferences drawn from Frederiksen's taxonomy. These she thought represented the kinds of inferences that primary aged students are most often asked to generate in response to classroom questions. The inferences are described and exemplified below with a sample statement that might appear in a reader's free recall protocol.

(a) attribute inference - generating an attribute not explicitly stated in the text.

Text: Susan's apology made Gail feel better.

Protocol: Susan's apology made Gail happy.

In this example, the reader has responded to the content words "feel better" and ascribed to Gail the attribute "happy".

(b) goal inference - generating a goal for an action.

Text: Cindy wiped up the spilled milk.

Protocol: Cindy wiped up the spilled milk so her mother would not get angry.

A goal not stated in text has been generated to supply a reason for the action of wiping up spilled milk.

(c) event inference - generating an event frame into which an object or proposition is placed.

Text: That is Michael. He is playing baseball. Crash!

Protocol: Michael broke the window.

This example illustrates the generation of an event frame into which a concept, Michael, was inserted in the role of agent.

(d) causal inference - the joining of two previously unconnected propositions in such a way that one is treated as the cause of the other.

Text: This is Sally. She is drying the dishes. Crash!

Protocol: While drying the dishes, Sally dropped and broke the glass.

In this example Sally's actions are interpreted to have caused the glass to break, thus connecting in recall two propositions that were previously separated.

Lipson's Research

An Overview of Lipson's Research

While investigating inferencing behaviour in children's reading, Lipson obtained results which indicated that certain inferences were more difficult to recognize than others. Children in grade three were asked to read aloud a total of eight passages. This was followed by a recognition test in which children responded to six paired recognition items constructed from information found in short expository passages they had just read. Each recognition pair consisted of two sentences, one of which represented a valid inference, whereas the other was an invalid inference. Validity was determined by whether the text contained the premise information upon which the inference was based. Without having an opportunity to refer back to the passage, children were asked to choose the best answer to each of the six paired items. Immediately following this task, the child was asked to recall all that could be remembered of the passage just read. This procedure was repeated for each of eight passages.

The recognition items were exemplars of four different types of inferences identified in Frederiksen's taxonomy of text-based inferences, namely, attribute, event, goal and causal inferences. The reader's comprehension of two types of explicitly stated information, explicitly stated attributes and explicitly stated events, also was assessed by recognition items paralleling the first and third types of

inferences.

Lipson found a reliable difference between performance on explicit and implicit items in that children recognized more explicit than implicit information. Of greater interest, however, was her finding that reliable differences were found among inferences. Attribute and goal inferences were more readily recognized than were event and causal inferences.

Lipson's Research: Its Strengths

First, she acknowledged that results from research using adult subjects may not be transferable to classroom settings where there is a need to understand how less mature readers comprehend information not explicitly stated in text. Secondly, her use of expository text reflects a needed departure from much of the reading research done with children which has usually used narrative material. Thirdly, her decision to focus attention on the general category of implicit or text-based inferences is a point which differentiates her work from that of many others investigating comprehension of unstated information. Until recently such inferences had been largely overlooked by psychologists (Johnson-Laird, 1983). They began to receive attention when computer programmers, attempting to design programs capable of understanding discourse, realized their ubiquity. The frequency with which they must be made speaks not only to the importance of their role in discourse comprehension, but also to the

need to address research attention to them. Finally, as was mentioned earlier, her decision to go beyond a dichotomous comparison of major types of inferences and examine performance on four different types of inferences is an important step forward. For these reasons, Lipson's study makes a positive contribution to inferencing research.

Critique of Lipson's Research

There are, however, aspects of Lipson's experimental design which leave certain questions unanswered -- questions that need answering before her results can be accepted with confidence. The detailed critique of Lipson's study presented now addresses those aspects of her methodology and interpretation which are weak, with one exception. A discussion of her results, which are considered a positive feature of her work, is included here as it provides the information necessary to develop the section addressing the interpretation of her findings. Topics to be discussed are subjects, passages, recognition items, procedures, results and interpretation. Each section includes a discussion of the questions her approach leaves unanswered and a rationale for modifying or leaving unchanged the features discussed.

Subjects

Lipson included only average and below average readers in her sample. She reported no rationale for this choice. A distinguishing

feature between skilled and unskilled readers is that the former display greater proficiency in comprehending implicit information (Oakhill, 1984; Waller, 1976; Wilson, 1979). As well, it has been found that not-so-fluent readers make inferences that tend to substitute general terms for specifics (Bridge, Tierney, & Cera, 1977, cited in Marshall, 1981), but fail to make inferences that integrate ideas (marshall, 1977, cited in Marshall, 1981). Apart from attribute inferences, the ability to relate ideas in text to one another is essential for the type of inferences examined here. For some reason goal inferences, though requiring the integration of ideas, appear to be processed competently. One might speculate that this finding is related to children's extensive exposure to narrative material which proceeds by the serial presentation of events and resolution of goal states. However, children do not fair as well on event and causal inferences which also depend on the construction of unstated relationships. The question that arises is whether Lipson's results can be generalized across all levels of reading proficiency, or whether they would remain isolated within the skill groups investigated in her study. Are event and causal inferences truly more difficult, or are they only found so by readers known to be less proficient at inferential comprehension? To answer this question the current research sampled skilled readers as well as those at skill levels represented in Lipson's study.

Passages

When the purpose of research is to examine the comprehension of expository text, it is advisable to stay within a single discipline, as the organization of information may vary across subject areas (Black, 1983). For instance, the arrangement of content in scientific writing may differ from that found in the behavioural sciences. In addition, the overall organization of text is known to have an effect on comprehension (Meyer, 1975; Meyer & Rice, 1984), hence the advice to restrict content to one subject area.

Lipson did not adhere to this recommendation. Her passages contain content from two areas of the curriculum: science and social studies. However, given the age level for which these materials were developed, this would not seem problematic as the passages would appear to fall into a text structure labelled "pseudo-narrative" by Gallagher and Pearson (1982). It is likened to narrative materials in that expository information is delivered by way of characters performing various actions (Englert & Hiebert, 1984). This pseudo-narrative structure is prevalent in elementary science and social studies texts. For this reason, Lipson's decision to include content from both social studies and science was left unchanged in the study conducted here.

Questions

Amongst recognition items designed to test comprehension of like inferences, there appeared to be differences in the type of reasoning required to select the correct answer. Goal inferences for the passages Eskimos and Venus Flytrap illustrate this point and are noted below. The first sentence in each pair represents the correct choice.

Eskimos often travel long distances to find food.

Eskimos often travel long distances to see friends.

Flies land on the Venus Flytrap to eat the sticky leaves.

Flies land on the Venus Flytrap to lay their eggs.

On the basis of information contained in the Eskimo passage both options are plausible as the passage discusses the need to travel long distances for both purposes. However, the context of the text deems the first the more appropriate answer. In the Venus Flytrap passage no mention is made of laying eggs, but information is provided about the attraction flies have for the plant's sweet-smelling sticky leaves. The conjecture that these items may require different types of reasoning was untested, leaving unanswered the possible effect these differences may have had on performance. Bearing this issue in mind, the decision was made nevertheless to use the recognition items

created by Lipson in replicating her study. To restructure them would be to stray too far from the original research, making it difficult, if not impossible, to compare results. It was reasoned that if such discrepancies caused the items to be meaningfully different, this would be detected during data analysis and would provide the basis for further research.

Procedure

In Lipson's study, the text was not available to readers during the recognition task. The question that arises is this: Could it be that poor performance on some types of inferences reflects the influence of memory rather than a lack of processing ability?

The empirical basis for this question comes from a study by Oakhill (1984) who investigated the relationship between inferencing and memory using narrative text. She found that the less skilled readers remained poorer at answering inferential questions even when the text was available. Oakhill concluded that her results provided no support for the claim that differences in ability to make inferences can be attributed to differences in memory for prose. Oakhill, however, did not differentiate between types of inferences. Therefore, her claim regarding the lack of an effect for memory may not apply to all types of inferences. Could it be that theoretically more difficult inferences are affected by memory in a manner unlike those considered less difficult? As well, can results obtained using

narrative material be generalized to expository text? In as far as the current study employed expository passages and incorporated a text-absent and text-present condition, it was hoped that answers to these questions would be provided.

Another question which merits consideration and which was addressed by extending Lipson's research is whether children who have difficulty recognizing or generating inferences are capable of identifying the premise information upon which that inference is based. The answer to this question would provide information useful for diagnostic and prescriptive purposes. Simply knowing that a child performs poorly on inference questions contributes little to understanding why this may be so or how the deficiency may be corrected. If it were to be found, for example, that inability to identify premise information impedes performance on inference questions, then instruction focusing directly on the deficient skill can be designed and initiated. These issues are pertinent given the current emphasis in reading research to <u>teach</u> comprehension (Durkin, 1978-79; Pearson, 1984). Identifying the source of a child's difficulty is the first step toward remediation. Accordingly, this study included a premise identification task. In the text present condition children were asked to underline the information used to determine what they considered the correct response.

Results

Lipson's results indicated that inference questions were significantly more difficult to recognize than those testing explicitly stated information, and that certain types of inferences were more difficult than others. The first finding would be expected as there is both empirical and theoretical support for the claim that comprehending explicit information is less demanding than making inferences. The second finding, however, is of greater interest because it provides information that many other studies on inferential processing do not. Specifically, it indicates that certain types of inferences are more difficult to process (Olson, 1984). Other investigators have found that implicit inter-sentence relations are more difficult for children (Bormuth, Carr, Manning, & Pearson, 1970; Robertson, 1968). However, the value of Lipson's findings lies in their degree of specificity and because they were obtained using expository text. For these reasons Lipson's research is considered somewhat unique and therefore needing replication before results can be accepted with confidence.

Interpretation of Results

In explaining her results, Lipson referred to Tulving and Thomson's (1973) notion of encoding specificity. According to Tulving and Thomson, what occurs during encoding is what predicts performance at time of retrieval. That is to say, if the cues presented in a

retrieval task do not coincide with aspects of a representation established during encoding, the probability of retrieving target information is reduced. Lipson argued that if a reader does not generate inferences at the time of reading, inference statements will not serve as effective retrieval cues. Her finding that inferences in general, and event and causal inferences in particular, were more difficult to recognize was interpreted as meaning they functioned poorly as retrieval cues.

This line of reasoning implies that inferences are generated only at the time of encoding. There is evidence to suggest, however, that inferences are constructed not only during encoding, but also at the time of retrieval (Kintsch, 1974; Schank, 1975), a point Lipson herself reports earlier in her paper. Recognizing that inferences can be constructed at the time of retrieval gives rise to the possibility that failure to recognize certain types of inferences may be associated with an inability to recall the premise information upon which that inference is based and not necessarily the inability to perform the cognitive processing demanded of the task. To control for this possible confound to Lipson's study, the current research had children respond to the recognition items under two conditions, text absent and text present.

Related to the issue of memory and the inability to recall premise information is the matter of saliency or relevancy of information tested (Warren, Nicholas, & Trabasso, 1979). The more a

piece of information contributes to the coherence of text, and thereby to overall understanding of the passage, the more a reader would be inclined to include it in recall (Crothers, 1979). Recall protocols, therefore, become a means of assessing the relevancy of information tested by recognition items. It is acknowledged that absence of targeted content from recall protocols is not necessarily an indication that it was not processed and therefore not available for inference generation. What its absence may suggest is that the information it conveyed was not particularly important to the overall understanding of the passage.

Focusing on the more difficult items, event and causal inferences, the question that emerges is whether children responded poorly to these questions because the content they tested was not particularly relevant. If this were true, it would not be surprising to find this information missing from recall protocols. If premise information were missing for the inferences Lipson labelled more difficult, caution would have to be exercised in interpreting children's poor performance. While poor performance could be attributed to processing difficulty, consideration would also have to be given to the possibility that inference generation relied on relatively unimportant information. This would lessen the likelihood of that information being processed as deeply as that which would contribute to overall passage comprehension. Examining recall protocols to determine if they contained premise information for the

targeted inferences would shed some light on this issue.

In the original study the recognition task preceded free recall. This sequencing is problematic because questions may serve as cues to memory, resulting in recall protocols containing information that might otherwise have been omitted. To preclude this possibility students in the present study were asked to recall each passage before presentation of the recognition task.

Summary

The above discussion has focused on those aspects of Lipson's research which were considered to present possible confounds. The research questions evolving from this discussion are summarized below:

1. Do skilled readers experience the same difficulty in recognizing event and causal inferences as did the participants in Lipson's study, or do the results describe a situation specific to the sample studied by Lipson?

2. What effect, if any, does memory have on children's ability to recognize the type of inferences investigated by Lipson?

3. Is it possible that children who score poorly on inference questions can, however, identify the premise information upon which those inferences are based?

4. Are event and causal inferences truly more difficult, or can the findings be explained by the irrelevancy of information tested? Irrelevant information would interfere with the depth of processing required to generate these inferences.

CHAPTER THREE

Subjects

Twenty-nine grade three students from two Coquitlam elementary schools participated in this study which was conducted during late May and early June. Participation was voluntary and required the permission of the students' parents. Of a possible thirty-three participants, nineteen girls and ten boys returned parental permission slips. Students ranged in age from 96 to 118 months. All were considered by their teachers to be competent users of English.

Teachers classified their students as reading below, at, or beyond grade level for both decoding and comprehension skill. The evaluation was based on standardized tests and classroom performance. In one school, the Gates MacGinitie, Form 1, Primary C was used. In the other, The District Levels Tests were used. Both teachers reported that formal testing confirmed their initial evaluation of student performance.

Only four children's teachers' rating on decoding and comprehension scores differed in that they were considered more skillful in one area than they were the other. Because assistance was provided for decoding difficulties, it was reasoned that deficiencies in this area would be minimized. Therefore, the decision was made to categorize students only according to their comprehension skill.

Materials

Passages

A preliminary study was conducted to determine the passage dependency of recognition items. It was conducted approximately six weeks prior to the experimental session and involved the same children who were to participate in the study. Students did not read the passages prior to answering the recognition questions Lipson had used. Instead, the experimenter read aloud each pair of recognition items. Students were instructed to listen and to mark which of the two sentences they considered correct. Children were advised that should they be unable to decide which of the sentences was correct, they should register their uncertainty by placing a question mark in the appropriate space. It was emphasized that this response was preferable to guessing. For each passage, presentation order of the items was determined by tossing a die. The arrangement of paired recognition items was decided by flipping a coin.

For three of Lipson's eight passages, the children's ability to answer the questions correctly without reading the passages exceeded chance. Therefore, Lipson's passages entitled Dinosaurs, Insects, and Totem Poles were eliminated. Further description of this pretest study is provided in Appendix B.

Time constraints required a fourth passage be eliminated. A maximum of sixty minutes was available for each subject, and while skilled readers could have accommodated more than four paragraphs in

that time, less skilled readers could not. Hence Lipson's passage entitled Whales was eliminated. The preliminary study had shown that of the five remaining passages considered acceptable, Whales contained content with which the children were most familiar.

The four passages selected for study then were Eskimos, Venus Flytrap, Vikings, and New Guinea. According to Lipson, the content of Eskimos and Vikings was considered similar to that found in grade three social studies and science curricula. The other two passages, Venus Flytrap and New Guinea, were included on the basis that they presented information new to the children. It will be remembered that Lipson was investigating children's performance on text-based inferences which are generated primarily from information available in text rather than prior knowledge. For this reason it was considered appropriate to use passages containing information about which the children were unlikely to have prior knowledge. Material for all the passages was drawn from commercial children's books, basal readers and social studies and science texts.

The passages were structurally similar on the following variables: number of words, number of sentences, and number of syllables. The number of words ranged from 174 to 177, the number of sentences from nineteen to twenty-one, and the number of syllables from 118 per 100 words to 128 per 100 words. While number of propositions was not mentioned by Lipson, it is reported here as it is considered a more accurate predictor of difficulty than number of

words or syllables (Kintsch, 1974). Using a method of propositional analysis developed by Bovair and Kieras (1981), Lipson's passages were found to be relatively similar in this regard. The number of propositions per passage ranged from 49 to 56.

Passages were typed on eight and half by eleven inch paper and were double spaced. They were packaged prior to the experimental session, with the sequence of passages in each package randomly arranged to control for order effects.

Recognition Item: Measures of Inferencing Ability

The recognition items developed by Lipson to measure inference making were used in this study. There were six types of items for each passage. Two items tested comprehension of explicitly stated information; four tested implicit information. For each paragraph, therefore, there were six recognition items. This resulted in a total of twenty-four items, with four exemplars of each type of item, one from each passage.

In each item, a correct statement was paired with a similar but incorrect statement. For example, a statement containing an attribute inference was paired with another statement also containing an attribute. What differentiated the two was that the former contained an inference generated from information available in the text while the other did not.

The twenty-four paired recognition items were used for both pre-testing and post-testing. Each recognition pair was typed double-spaced on a white index card measuring six by four inches with approximately one and a half inches separating the two options. The toss of a coin determined which of the two items appeared first. All questions for the same story were mounted on similarly coloured backing and then laminated.

Procedures

Students were seen individually in a quiet location away from their regular classroom. They were first asked to print their names on the data collection sheets. They then received the following instructions:

> We're going to be reading some stories together. I would like you to read them carefully because I'll be asking you to tell each story back to me. I'll also be asking you some questions about the stories we've read.

Oral Reading

Having confirmed that participants had a general understanding of what was to follow, the first passage was presented. Children were asked to read the passage aloud. They were advised that help would be provided for any words they found difficult to read.

Recall

Having read the passage, the sheet upon which it was printed was turned face down. Each child was then asked to tell the story back to the experimenter. A few children queried the meaning of these instructions, asking if they were to retell the story as it appeared on the page. This question received the following response:

> Tell the story back to me the best way you can. If you can't remember the exact words used, tell me using your own.

Recall protocols were recorded on an audiotape cassette for later transcription. When each participant had finished, a prompt was delivered:

> Is there anything else you'd like to tell me?
> Is there anything you remember that you haven't already mentioned?

Upon being assured that nothing further was forthcoming, the next task was introduced.

Recognition Task - Text Absent

Participants were then asked to respond to the six paired recognition items in the text absent condition. The cards on which the items were typed were shuffled to ensure random presentation. Cards were presented individually and placed on the table in front of the child. The experimenter pointed to each sentence while reading it aloud. The child was asked to indicate which of the two sentences was correct. The response was recorded by the experimenter on a separate sheet of paper.

Recognition Task - Text Present and Premise Identification

After completing the six questions, the sheet on which the passage was printed was re-presented to the student. The same recognition items were administered again in the same order. Participants were advised to consult the passage before providing an answer. Participants also were instructed to underline the words or the sentences they used in selecting the correct statement. Each type of recognition item (e.g., event inference) had been colour coded, and students used an appropriately coloured pencil crayon to underline the information they used to answer each item.

Before proceeding with this task, subjects were asked whether the instructions had been understood. Children's performance was monitored carefully and further clarification was provided when it appeared necessary. This caution was taken because children of this

age (a) may be reticent to acknowledge lack of comprehension; (b) may not realize they do not understand what is to be done until they are into the task; or (c) may not ask for clarification because they place the responsibility for misunderstanding with themselves and not with the instructions (Markman, 1977). For the few children who did require further assistance, repeating the instructions was usually sufficient. When it was not, the child was instructed to reread the first sentence in the passage, and then asked if it contained information which could be used to answer the question. This procedure, repeated two or three times, provided the necessary clarification.

Upon completing this task, the experimenter engaged the child in casual conversation and then began the same sequence of tasks with the next passage. Before beginning subsequent tasks for each of the remaining passages, a shortened version of the instructions noted above was delivered.

Concluding The Session

Upon completing all tasks for all passages, the students were thanked for their participation and encouraged not to repeat the questions to their classmates.

Scoring Procedures

Recognition Task

In both the text present and absent conditions, the recognition items were scored either right or wrong, one point being awarded for each correct response.

Recall Protocols

Recall protocols transcribed from the audiotapes were scored based on Bovair and Kieras' (1981) adaptation of Kintsch's (1974) system for propositional analysis. First, each stimulus passage was propositionalized. Sentences in a student's protocol were then examined to determine what propositions from the original text they contained. A point was scored for each proposition recalled in either verbatim or semantically equivalent form. Semantically equivalent form is illustrated in these two examples, and is defined as those situations in which the words may differ, but not the meaning.

Text: The must share their food.

Recall: They have to share their food.

Text: They build new igloos.

Recall: Again they make igloos.

A list of other semantic equivalencies occurring in the protocols is reported in Appendix C. A half point was given for incomplete propositions, namely, a proposition from which an essential element was either missing or labelled incorrectly. These examples illustrate the procedure.

Text: "Wig men" wear big headdresses made of human hair.

Regarding the first proposition, a student's sentence would receive a half point for mentioning "headdresses", but lose a half point if it failed to specify "wig men". Similarly, in the second proposition, a student's sentence would receive a half point for mentioning that headdresses were made of something "human" but forfeit a half point for mistakenly referring to that substance as "fur" instead of "hair".

A half point was also scored when it appeared that a recalled proposition captures an <u>essential</u> aspect of a stimulus proposition, if not its <u>exact</u> meaning. An example follows:

Text: Once a year...

Recall: Every spring...

Identification of Premise Information

As previously explained, children were asked to identify the words or sentences which enabled them to respond to the recognition items. A point was scored each time a child's underlining corresponded with the premise information identified by the procedures described in Chapter Four.

Inclusion of Premise Information in Recall

Recall protocols were also analyzed to determine whether they included the premise information for the inferences tested. A dichotomous scoring procedure was used to record their presence or absence.

CHAPTER FOUR

Data Analysis and Results

The purpose of this study was to determine the replicability of Lipson's results and to examine the possible effect of memory and saliency of information that was tested on children's performance. The order in which these issues are examined is: the effects of type of information, type of inference, memory, the ability to identify premise information and the recall of premise propositions. This study also examined the effect of reading level for each of the above factors. The chapter concludes with a comparative analysis of different measures of comprehension.

In the analyses that follow, results are reported that are statistically reliable at the .90 level or better. This is contrary to the tradition that arbitrarily limits discussion to results that are statistically reliable at the .95 level or better. It was reasoned that the more liberal range of significance was merited for the following reasons. One, the current research, though in part a replication study, was also exploratory in nature. There was, therefore, an interest in identifying trends in the data. Two, the overall sample size was small. This becomes more problematic when the data is subdivided into and analyzed by reading level. A further complication was variation in cell size associated with the teachers' ratings of comprehension (11, 6, 11 for those reading below, at, and

above grade level respectively). It may be argued that this would be reason to place more stringent confines on the level of statistical reliability. However, the more liberal criterion was selected to minimize the possibility of Type II error.

Finally, while some types of questions were answered more accurately than others, and while some students performed at a higher level of competence than others, the scores were generally very high. This necessitates considering a ceiling effect and approaching with caution the interpretation of data associated with non-normal distributions. Under these conditions it was reasoned that decreasing the level of statistical reliability to .90 would once again lessen the likelihood of Type II error.

A further note regarding statistical reliability is that all alpha levels at .01 or less are reported as .01.

One female student was eliminated from the analysis as no recall protocol could be elicited. Conversations with the classroom teacher indicated that this child's performance fell outside the normal range of accomplishment, as defined by the level of achievement of other children in the class.

Analysis of Type of Information Recognized

Following elimination of one student, 28 participants remained.

A 3 (reading level) x 2(type of information) x 2 (text absent vs. present) x 4(passages) multivariate analysis of variance was

performed. In this analysis a priori contrasts were conducted to test for differences between composite scores for attribute and goal inferences and event and causal inferences. Lipson's results provide the empirical basis for the decision to run a priori contrasts. The theoretical foundation is provided by Frederiksen's taxonomy which gave rise to the hypothesis that the larger the semantic unit needed to generate the inference, the more difficult the inference is to construct.

In running this MANOVA and the one which follows, it was decided to test first the overall effect of type of information (explicit vs. implicit) and type of inference (the four levels of implicit information), regardless of text-absence or text-presence. In the event that Lipson's results were replicated, it was considered important to determine whether the effects generalized over presentation condition. Should the statistics warrant it, presentation condition would be dissected thrcugh further analysis. All a posteriori contrasts were tested by conducting Scheffe post hoc tests.

Scores for the two explicit items and the four implicit items served as the dependent variables. Reading level was the sole between subjects factor. Passages, type of information and presentation condition were within-subject factors. Descriptive data appear in Table 3.

Table 3

The Mean Percentage of Correct Answers
as a Function of Type of Information and Text Absence or Presence

	Eskimo		New Guinea		Venus Flytrap		Vikings	
	Expl.	Impl.	Expl.	Impl.	Expl.	Impl.	Expl.	Impl.
No Text	.86 (.24)*	.73 (.21)	.86 (.32)	.71 (.22)	.91 (.20)	.71 (.29)	.96 (.15)	.66 (.23)
Text	.82 (.25)	.71 (.25)	.91 (.20)	.73 (.18)	1.00 (.00)	.68 (.28)	.91 (.20)	.64 (.21)
No Text	.83 (.26)	.50 (.22)	1.00 (.00)	.88 (.21)	1.00 (.00)	.75 (.16)	1.00 (.00)	.79 (.25)
Text	.92 (.20)	.58 (.13)	1.00 (.00)	.86 (.14)	.92 (.20)	.79 (.19)	1.00 (.00)	.79 (.10)
No Text	.96 (.15)	.75 (.19)	.96 (.15)	.77 (.18)	.96 (.15)	.84 (.17)	1.00 (.00)	.75 (.16)
Text	1.00 (.00)	.77 (.21)	1.00 (.00)	.80 (.25)	1.00 (.00)	.91 (.13)	1.00 (.00)	.91 (.13)
No Text	.89 (.21)	.69 (.22)	.93 (.22)	.77 (.20)	.95 (.16)	.77 (.23)	.98 (.10)	.72 (.21)
Text	.91 (.20)	.71 (.22)	.96 (.13)	.79 (.20)	.98 (.09)	.80 (.23)	.96 (.13)	.78 (.20)

*Standard deviations appear in brackets.

Main effects. There were no statistically reliable differences on the within subjects factor of passages, however a reliable main effect was found for each of the other factors. Students reading above grade level reliably outperformed all others, and those reading at grade level excelled over those reading below ($F(2,25) = 6.21$, $p < .01$). Finally, though the effect was only marginally reliable, performance was slightly superior in the text present condition ($F(1,25) = 3.39$, $p < .08$).

Interactions. A test for interactions between reading level and passages, presentation condition, and type of information showed no reliably different results. Nor were any statistically reliable interactions found between passages, presentation condition, and type of information.

Analysis of Inference Types

To determine whether inference type affected performance a 3 (reading level) x 4(inference type) x 4(passage) x 2 (text absent versus present) multivariate analysis of variance was performed. In this analysis the scores for each of the sixteen inference items were employed as the dependent variables. Means and standard deviations are presented in Table 4.

Table 4

The Mean Percentages of Correct Answers as a Function of Type of Inference and Text Absence or Presence

Reading Level		Eskimos				New Guinea				Venus Flytrap				Vikings			
		Att.	Goal	Event	Causal	Att.	Goal	Event	Causal	Att.	Goal	Event	Causal	Att.	Goal	Event	Causal
Below	No Text	.36 (.51)*	1.00 (.00)	1.00 (.00)	.55 (.52)	1.00 (.00)	.91 (.30)	.46 (.52)	.46 (.52)	.73 (.47)	.91 (.30)	.55 (.52)	.64 (.51)	.64 (.51)	.91 (.30)	.64 (.51)	.46 (.52)
	Text	.46 (.52)	1.00 (.00)	1.00 (.00)	.36 (.51)	.91 (.30)	.91 (.30)	.55 (.52)	.55 (.52)	.73 (.47)	.91 (.30)	.55 (.52)	.55 (.52)	.64 (.51)	.91 (.30)	.55 (.52)	.46 (.52)
At	No Text	.00 (.00)	1.00 (.00)	.67 (.52)	.33 (.52)	.63 (.41)	1.00 (.00)	1.00 (.00)	.67 (.52)	.50 (.55)	1.00 (.00)	.67 (.52)	.83 (.41)	.83 (.41)	1.00 (.00)	.83 (.41)	.50 (.55)
	Text	.17 (.41)	1.00 (.00)	.67 (.52)	.50 (.55)	1.00 (.00)	1.00 (.00)	.83 (.41)	.67 (.52)	1.00 (.00)	1.00 (.00)	.67 (.52)	.50 (.55)	1.00 (.00)	1.00 (.00)	1.00 (.00)	.17 (.41)
Above	No Text	.46 (.52)	1.00 (.00)	.73 (.46)	.82 (.41)	1.00 (.00)	.91 (.30)	.36 (.51)	.82 (.41)	1.00 (.00)	1.00 (.00)	.64 (.51)	.73 (.47)	.91 (.30)	1.00 (.00)	.64 (.51)	.46 (.52)
	Text	.64 (.51)	1.00 (.00)	.82 (.41)	.64 (.51)	.91 (.30)	1.00 (.00)	.55 (.52)	.73 (.47)	1.00 (.00)	1.00 (.00)	.82 (.41)	.82 (.41)	1.00 (.00)	1.00 (.00)	1.00 (.00)	.64 (.51)
All	No Text	.32 (.48)	1.00 (.00)	.82 (.39)	.61 (.50)	.96 (.19)	.26 (.93)	.54 (.51)	.64 (.49)	.79 (.42)	.96 (.19)	.61 (.50)	.71 (.46)	.79 (.42)	.96 (.19)	.68 (.48)	.46 (.51)
	Text	.46 (.51)	1.00 (.00)	.86 (.36)	.50 (.51)	.93 (.26)	.96 (.19)	.61 (.50)	.64 (.49)	.89 (.32)	.96 (.19)	.68 (.48)	.64 (.49)	.86 (.36)	.96 (.19)	.82 (.39)	.46 (.51)

* Standard deviations appear in brackets.

Main effects. This analysis revealed no main effect for either passages or presentation condition. However, statistically reliable differences were found among reading levels ($F(2,25) = 4.61$, $p < .02$) and types of inferences ($F(3,23) = 124.45$, $p < .01$). Children reading above grade level outperformed all others, and participants reading at grade level fared better than those reading below. A priori F-tests were conducted to interpret the effect of inference type. The decision to use a priori tests was based on the hypothesis that event and causal inferences would be more difficult to recognize as they were constructed from larger semantic units than either attribute or goal inferences. It was found that a composite score for attribute and goal inferences was higher than a composite score for event and causal inferences ($F(1,25) = 11.34$, $p < .01$). Within the broad categories of "easier" and "harder" inferences, there were no reliable differences between attribute and goal inferences, while there were between event and causal items. Event inferences were more easily recognized than causal inferences ($F(1,25) = 197.46$, $p < .01$).

Interactions. Reading level interacted reliably with inference type ($F(6,46) = 3.48$, $p < .01$); marginally with passages ($F(6,46) = 2.02$, $p < .06$); and not at all with presentation condition. An interaction was also found between passages and inference types ($F(9,17) = 5.16$, $p < .01$). The final interaction was that between inference type and presentation condition ($F(3,23) = 6.72$, $p < .01$).

For reasons not understood, overall performance on causal questions declined in the text present condition.

Analysis of the Effect of Memory

The two MANOVAS just described also tested for the possible effects of memory. As was described earlier, children answered questions under two conditions: text-absent and text-present. It was reasoned that differences in scores would be a means of assessing the affect of memory on performance. As previously reported, presentation condition was found to have only a marginal effect on the ability to respond to different types of information ($F(1,25) = 3.39$, $p < .06$), and none of the ability to respond to different types of inferences. In neither MANOVA did presentation condition interact significantly with reading level, and only in the second that investigated the effect of types of inferences did it interact reliably with question type ($F(3,23) = 6.72$, $p < .01$). Unexpectedly it was found that, overall, performance on causal inferences declined when the text was available.

Analysis of the Ability to Identify Premise Information

In order to better understand at what point in cognitive processing inferential reasoning breaks down, it was decided to examine the participants' ability to identify the information in text upon which each of the recognition items was based. The question of

interest was if children who had difficulty responding to a recognition item could identify the premise information from which the inference could be generated.

Before identification of premise information could be scored, it was considered necessary to develop a list of propositions constituting acceptable premise information. This listing was to be generated by asking university students, assumed to be skilled readers, to perform the same task as the grade three students. The undergraduate participants were divided into four groups of approximately thirteen each. Each group was assigned one of the four stimulus passages used in the study. This was read prior to the presentation of the six paired recognition items. Participants were to indicate which of the two options they considered correct and then underline the information in the passage upon which they based their decision.

It was first thought that the desired list would include those words and sentences underlined by at least sixty percent of the adult readers. There were difficulties, however, in using consensus to develop this list. To rely solely on consensus would be to render invalid logically acceptable information, simply because it had not been selected by an arbitrarily established minimum number. It was noted as well that in at least one case there was a discrepancy between logically valid content words selected by adults and those chosen by children. The information identified by grade three

students, which was reasonable and provided sound basis from which to generate the targeted inference, was not that chosen by the majority of adult readers.

There was a third difficulty. Some adult readers indicated that as many as three sentences were used for decision-making purposes. Seldom, however, were children as inclusive. In studying this difference, it was concluded that the more extensive adult underlining included not only critical information, but also supportive statements which were not essential but increased the certainty with which the response could be made.

It was decided to reject the notion that consensus agreement was required. Instead, any information identified by adult readers which appeared logically correct would be considered as legitimate support for the correctness of the student's choice.

There was the odd occasion when a child selected words that no adult had underlined. Here an attempt was made to relate this information to that which had been identified by the adult readers. This was done by tracing it back through the propositional analysis of the passages. In most cases this was successful. When it was not, a judgement was made as to whether it seemed reasonable that the targeted inference could be generated on the basis of the child's selection. A listing of such decisions appears in Appendix C.

In some cases a child underlined a single word. This was considered correct if that word formed part of the proposition

selected by adult readers. The one stipulation was that the solitary word must reflect a critical part of the identified proposition. That is to say, it must contain information which determined the correctness of the response. It could not be a predicate or argument common to both options in a test item. An example will demonstrate this distinction.

(a) Sled dogs are strong and smart.

(b) Sled dogs are strong, but not very smart.

In sentence (a) the word "strong" is common to both sentences and is not considered to be the information which differentiates the options. On the other hand, if the solitary word "smart" in the first sentence were identified, it would be adjudicated as having provided sufficient support for the selection of sentence (a). It was reasoned that the words "smart" and "not very smart" constitute the critical content which discriminates one sentence from the other.

These results for scoring information were applied to the propositionalized stimulus materials and recall protocols. The procedure provided a means of determining what premise information had been identified.

A two-step procedure was used in analyzing the resulting scores. First, it was necessary to determine not only how well children from each reading level were able to recognize pertinent premise

information but also what effect question type had on performance. Descriptive statistics were then used in the second part of the analysis, the calculation of conditional probabilities. The purpose of these calculations was to determine the likelihood of premise information being identified when the recognition item was answered incorrectly. A discussion of the MANOVA will be followed by a description of the conditional probabilities.

A 3(reading level) x 6(question type) x 4(passage) multivariate analysis of variance was performed to assess the effects of reading level and the ability to identify premise information on response to recognition items. All means and standard deviations are reported in Table 5.

Main effects. The analysis revealed main effects for reading level ($F(2,25) = 5.77$, $p < .01$) and question type ($F(4,22) = 16.27$, $p < .01$), but none for passages. Overall, the performance of children reading above grade level reliably exceeded that of all other participants, and the performance of those reading at grade level was reliably superior to those reading below. Regarding question type, a priori contrast showed that premise information for explicit questions was identified more readily than that for implicit questions ($F(1,25) = 49.73$, $p < .01$). Within inferences, an a priori contrast revealed that more premise information was correctly identified for attribute and goal inferences than for event and causal inferences ($F(1,25) =$

Table 5

Mean Percentages of Correct Responses
to the Premise Identification Task
as a Function of Reading Level and Passages

Question	Reading Level	Eskimo		New Guinea		Venus Flytrap		Vikings	
		X	S.D.	X	S.D.	X	S.D.	X	S.D.
Explicit Attribute	Below	1.00	.00	.91	.30	1.00	.00	1.00	.00
	At	.83	.41	.83	.41	1.00	.00	1.00	.00
	Above	1.00	.00	.91	.30	1.00	.00	1.00	.00
	All	.96	.19	.89	.32	1.00	.00	1.00	.00
Explicit Event	Below	.72	.47	1.00	.00	.82	.41	1.00	.00
	At	1.00	.00	1.00	.00	1.00	.00	1.00	.00
	Above	.91	.30	1.00	.00	1.00	.00	1.00	.00
	All	.86	.37	1.00	.00	.93	.26	1.00	.00
Implicit Attribute	Below	.64	.51	.82	.41	.82	.41	.82	.41
	At	1.00	.00	1.00	.00	1.00	.00	1.00	.00
	Above	.91	.30	1.00	.00	1.00	.00	.91	.30
	All	.82	.39	.93	.26	.93	.26	.89	.32
Implicit Goal	Below	.64	.51	1.00	.00	.55	.52	.91	.30
	At	.33	.52	1.00	.00	.67	.52	1.00	.00
	Above	.55	.52	1.00	.00	.91	.30	1.00	.00
	All	.54	.51	1.00	.00	.71	.46	.96	.19
Implicit Event	Below	.55	.52	.55	.52	.46	.52	.64	.51
	At	.67	.52	1.00	.00	.67	.52	1.00	.00
	Above	.73	.47	.64	.51	.82	.41	.91	.30
	All	.64	.49	.68	.48	.64	.49	.82	.39
Implicit Causal	Below	.64	.51	.46	.52	.73	.47	.36	.51
	At	1.00	.00	.50	.55	.67	.52	.33	.52
	Above	.91	.30	.45	.52	.73	.47	.73	.47
	All	.82	.39	.46	.51	.71	.46	.50	.51

23.60, $p < .01$). There was a marginal difference for the planned contrast comparing attribute and goal inferences, with performance on the former being somewhat superior ($F(1,25) = 3.52$, $p < .07$). No reliable differences were found between event and causal inferences.

Interactions. Reading level was found to interact significantly with question type ($F(8,44) = 2.60$, $p < .02$). The performance of children reading at grade level on items testing comprehension of attribute and goal inferences appears responsible for this interaction. For a reason not understood, though one possibly related to sampling error, children reading at grade level outperformed all others on the identification of premise information supporting attribute inferences. Their performance on goal inferences, however, declined sharply. A plausible explanation is that small sample size may be responsible for this finding. Apart from the interaction just described, reading level did not interact reliably with any other factor.

The second step in this stage of the analysis involved calculating the probability of subjects not identifying the premise information when a recognition item had been answered incorrectly. This gauges the extent to which breakdown in inferential processing is related to an inability to identify relevant premise information. A tally was made of all questions incorrectly answered in the text present condition, but for which premise information had been

accurately identified. As can be seen from Table 6, the number of cases meeting these criteria are so few as to render meaningless the results obtained when these figures are divided by the total number of wrong answers for each type of inference.

Analysis of the Saliency of Information Tested

The next research question to be discussed is that assessing the role of saliency of information tested. Could it be that inferences Lipson found to be more difficult only appeared so because they were based on information to which children paid little attention and therefore had not processed deeply? As may be recalled, protocols were examined to determine if the propositions upon which each recognition item was based had been mentioned in recall. If they had, a point was scored. These points were summed for each of the 24 questions. The resulting figures were used as the dependent variables in a 3(reading level) x 6(question type) x 4(passages) multivariate analysis of variance. All means and standard deviations are reported in Table 7.

Main effects. No main effect was found for reading level. However the analysis did reveal statistically reliable differences for both passages ($F(3,23) = 17.76$, $p < .01$) and question type ($F(4,22) = 38.12$, $p < .01$). To interpret the results of the analysis of variance, univariate F-tests were applied a posteriori. Significantly

Table 6

Probability of Correctly Identifying Premise Information when Recognition Item is Answered Incorrectly

Questions	Eskimo			New Guinea			Venus Flytrap			Vikings		
	A	B	C	A	B	C	A	B	C	A	B	C
Impl. Attribute	16	19	.84	1	2	.50	5	6	.83	3	5	.60
Impl. Goal		*		2	2	1.00	1	1	1.00	1	1	1.00
Impl. Event	2	5	.40	6	13	.46	3	11	.27	8	9	.89
Impl. Causalute	9	11	.82	2	10	.20	3	8	.38	5	15	.33

A = the number of instances when a question was answered incorrectly but the premise information was accurately identified.

B = the number of times the question was answered incorrectly.

C = probablility of correctly identifying premise information when recognition item is answered incorrectly.

Table 7

The Mean Percentages of Times Recall Contained Premise Information as a Function of Reading Level, Question Type and Passages

Questions	Reading Level	Eskimo		New Guinea		Venus Flytrap		Vikings	
		X	S.D.	X	S.D.	X	S.D.	X	S.D.
Explicit Attribute	Below	.00	.00	.27	.47	.64	.51	.46	.52
	At	.17	.41	.67	.52	.67	.52	.83	.41
	Above	.18	.41	.55	.52	.64	.51	.64	.51
	All	.11	.32	.46	.51	.64	.49	.61	.50
Explicit Event	Below	.82	.41	.09	.30	.82	.41	.36	.51
	At	.17	.41	.00	.00	1.00	.00	.83	.41
	Above	.64	.51	.91	.30	.73	.47	.64	.51
	All	.61	.50	.07	.26	.82	.39	.57	.50
Implicit Attribute	Below	.73	.47	.64	.51	1.00	.00	.73	.47
	At	.33	.52	1.00	.00	1.00	.00	1.00	.00
	Above	.64	.51	.73	.47	1.00	.00	.91	.30
	All	.61	.50	.75	.44	1.00	.00	.86	.36
Implicit Goal	Below	.27	.47	.27	.47	.09	.30	.46	.52
	At	.67	.52	.67	.52	.83	.52	.67	.52
	Above	.55	.52	.64	.51	.46	.52	.73	.47
	All	.46	.51	.50	.51	.29	.46	.61	.50
Implicit Event	Below	.36	.51	.09	.30	.64	.51	.27	.46
	At	.50	.55	.00	.00	.50	.55	.50	.55
	Above	.27	.47	.09	.30	.55	.52	.46	.52
	All	.36	.49	.07	.26	.57	.50	.39	.50
Implicit Causal	Below	.09	.30	.00	.00	.55	.52	.00	.00
	At	.33	.52	.17	.41	.50	.55	.17	.41
	Above	.36	.51	.00	.00	.36	.51	.36	.51
	All	.25	.44	.04	.19	.46	.51	.18	.39

more premise propositions were recalled for the Eskimo passage than for New Guinea ($F(1,25) = 10.33$, $p < .01$). Similarly the number of recalled premise propositions for Venus Flytrap exceeded that of Vikings ($F(1,25) = 16.34$, $p < .01$). Finally, the number of premise propositions recalled for Flytrap plus Vikings was greater than that for Eskimos plus New Guinea ($F(1,25) = 24.58$, $p < .01$).

Regarding the main effect of question type, it was discovered that more premise propositions were recalled for explicit than implicit items ($F(1,25) = 13.98$, $p < .01$) and for attribute and goal inferences that event and causal ($F(1,25) = 22.62$, $p < .01$). It was also noted that premise propositions were recalled more frequently for attribute than goal inferences ($F(1,25) = 156.95$), $p < .01$), though no significant differences were detected for the number of premise propositions recalled for event and causal inferences.

Interactions. Apart from the marginally reliable interaction with passages ($F(6,46) = 1.89$, $p < .01$), for which the group reading at grade level is once again responsible, no other factors were found to interact with reading level. Interactions were discovered among passages and type of information ($F(3,23) = 11.59$, $p < .01$) and passages and level of inference question ($F(9,17) = 6.70$, $p < .01$). As can be seen from Figure 1, the interactions are complicated and difficult to explain. A general discussion regarding these findings is included in Chapter Five.

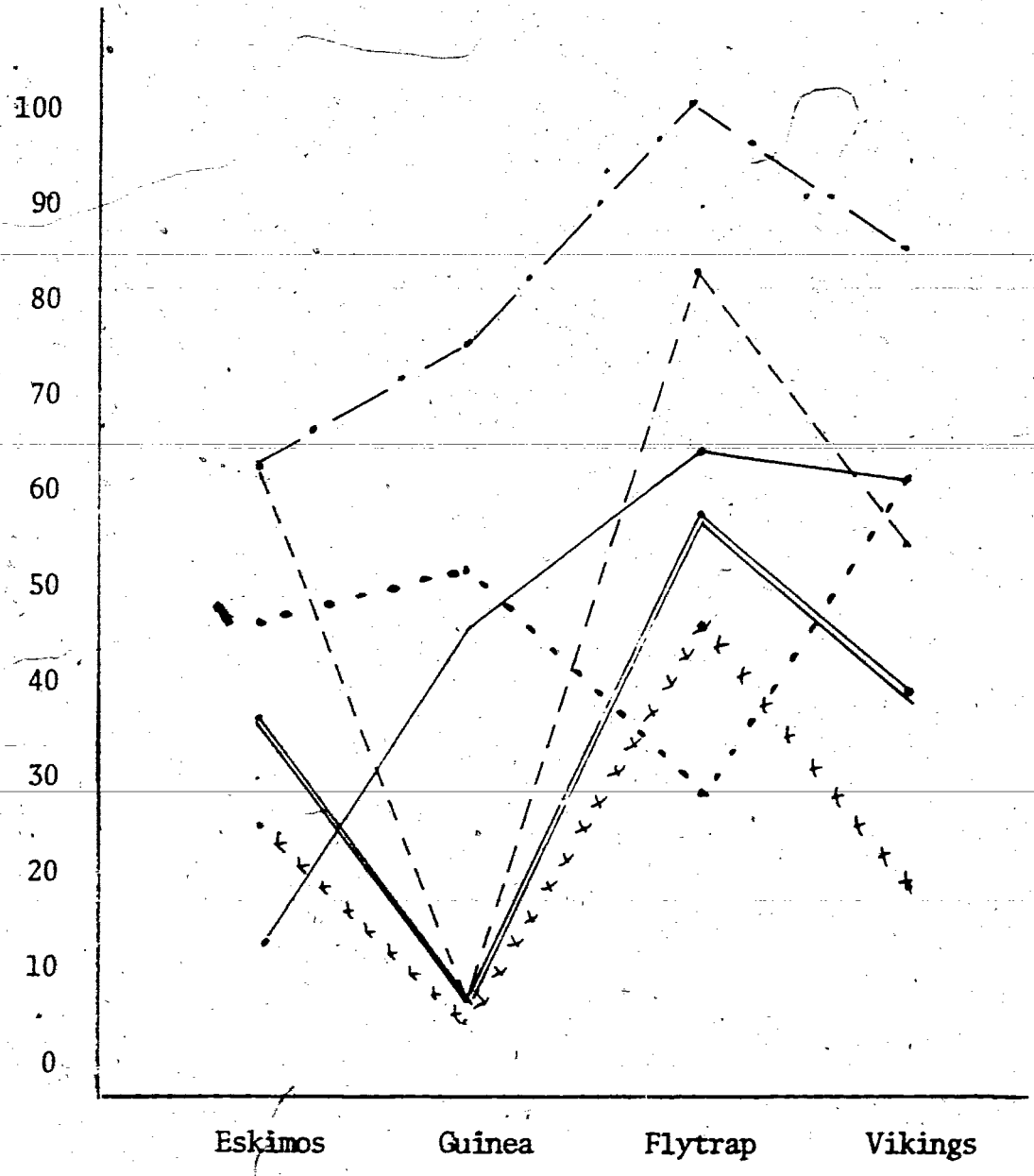


Figure 1. Interaction between question type and passages on the recall of premise propositions.

Explicit Attribute ————
Explicit Event _ _ _ _ _
Implicit Attribute —.—.
Implicit Goal . . . . .
Implicit Event ═══
Implicit Causal x x x x x

Also examined was the role recall of premise information played in answering recognition items that call for inferences. The first question asked was whether readers had to recall the premise information in order to answer a recognition item correctly. The average likelihood of a reader correctly answering an inference-based recognition item without recalling premise information was .51. This is indicative of chance-level performance for two-choice recognition items. The answer to the question noted above, therefore, is no. There is a pattern, however, in the likelihood of inferences being correctly recognized when premise information is not recalled. This pattern reflects the increasing order of theoretical difficulty hypothesized by Frederiksen in his taxonomy of text based inferences. Means were .16 for attribute inferences, .52 for goal inferences, .59 for event inferences, and .76 for causal inferences (Table 8). When compared to chance-level performance, it is statistically unlikely ($p < .05$) that recognition of an event inference did not depend on recalling the premises underlying this type of inference. A second question, do readers who recall premise information then answer the inference-based recognition items incorrectly, is responded to positively. It must be noted though that the likelihoods here show wide variance across items and passages and are based on few cases. The data, therefore, are not stable and merit very cautious interpretation. The overall likelihood of incorrectly answering an inference-based recognition item given recall of premise information

Table 8

The Probability of Correctly Answering the Recognition Item When the Premise Information Was Not Recalled

Implicit Question	Eskimo	New Guinea	Venus Flytrap	Vikings	Mean
Attribute	.22	.25	.00	.17	.16
Goal	.54	.50	.70	.33	.52
Event	.56	.87	.35	.58	.59
Causal	.76	.94	.55	.77	.76
Mean	.52	.64	.44	.46	.51

was .35, reliably different from chance-level performance ($p < .05$). Further analysis is not warranted given the characteristics of the data just mentioned.

Overall there are some differences in readers' abilities to "comprehend" text (measured by recognizing valid inferences) as a function of the type of inference called for and the recall of premise information needed to recognize the inference.

Analysis of Measures of Comprehension

In this study three different measures of comprehension were employed; recall of information following the reading of a text (Table 9); the ability to answer questions about information explicitly present in text; and the ability to answer questions that call for inferences based on premise information explicitly presented in a text. To explore the relation among these three different definitions of comprehension, correlations were computed among readers' scores on: one, each type of recognition item summed over the four passages; two, recall of propositions from each passage; and three, correctly answered recognition items summed over the six items within each passage (Table 10).

Comparisons across the definitions of comprehension showed strong discriminant validity. The median correlation among scores on the six types of items and the number of propositions recalled form each passage was .11 ($r = -.04$ to .43). For correlations among the six

Table 9

Average Number of Propositions Recalled Expressed as a Percentage

Reading Level	Eskimo		New Guinea		Flytrap		Vikings	
	X	S.D.	X	S.D.	X	S.D.	X	S.D.
Below	.26	.12	.15	.08	.22	.03	.21	.09
At	.36	.07	.29	.12	.28	.08	.28	.12
Above	.31	.18	.31	.20	.30	.13	.30	.17
All	.30	.14	.24	.16	.26	.10	.26	.13

Table 10

Correlations Among Readers' Scores on Each Type of Recognition Item Summed Over the Four Passages, Recall of Propositions from Each Passage, and Correctly Answered Recognition Items Summed over the Six Items Within Each Passage

		Q1	Q2	Q3	Q4	Q5	Q6	Total Proposition Recall E	G	F	V	No-Text Questions E	G	F	V	Proportional Proposition Recall E	G	F	V
	Q1		.60	-.23	.13	.35	.03	.10	.16	.38	.16	.12	.64	-.01	.18	.10	.16	.38	.16
	Q2			-.04	.27	.02	.20	.11	.19	.43	.01	.20	.51	.29	.11	.11	.19	.43	.01
	Q3				.06	-.12	.15	.04	.28	.07	-.04	.38	-.06	.21	.15	.03	.27	.07	-.04
	Q4					.45	.30	.12	.34	.13	.16	.12	.55	.31	.42	.12	.34	.13	.16
	Q5						-.28	.18	.26	.06	.28	-.02	.56	.35	.11	.18	.26	.05	.28
	Q6							.06	.15	.25	-.03	.29	.10	.26	.40	.06	.15	.25	-.03
Total Proposition Recall	E								.60	.62	.62	.14	.20	.06	-.02		.60	.62	.63
	G									.60	.63	.32	.41	.29	-.16			.60	.63
	F										.70	-.13	.28	.36	-.11				.70
	V											.20	.13	.19	-.19				
No Text	E Ques. sum												-.03	-.10	-.17	.14	.31	.25	.20
	G Ques. sum													.17	.19	.20	.41	.27	.13
	F Ques. sum														-.01	.06	.29	.36	.18
	V Ques. sum															-.00	-.15	-.11	-.19
Proportional Proposition Recall	E																.60	.62	.62
	G																	.60	.63
	F																		.70
	V																		

Q1 = explicit attribute; Q2 = explicit event; Q3 = implicit attribute; Q4 = implicit goal; Q5 = implicit event; Q6 = implicit causal.

E = Eskimo; G = New Guinea; F = Venus Flytrap; V = Vikings.

types of recognition items and scores on the recognition task for each passage, the median correlation was .24 (r = -.02 to .64). These are low despite being part-whole correlations. The median correlation between recall of passages and scores on the 6-item recognition tests was .16 (r = -.19 to .41).

CHAPTER FIVE

The first purpose of this study was to determine the replicability of Lipson's results. As Lipson found, and as would be expected from previous research investigating differences between implicit and explicit comprehension, children fared better on explicit than implicit items. As well, the current data confirmed Lipson's findings that, among different types of inferences, children had greater difficulty recognizing event and causal inferences than attribute and goal inferences. These findings held even with the inclusion of skilled readers. Exclusion of this group of readers from Lipson's study had led to uncertainty regarding the generalizability of her results across all levels of reading ability. The question arose of whether event and causal inferences were difficult for all readers, or whether this effect was confined to less skilled readers who are known to have greater difficulty with inferential processing of written discourse (Oakhill, 1984; Waller, 1976).

Having confirmed that Lipson's results were more generalizable than the original study might suggest; the next question asked was whether her findings could be related to the effect of memory. Could it be that the difficulty encountered on some items was associated with an inability to recall the information needed to generate the inference? If this were so, having the text available during question answering should remove effects attributable to deficiency in memory.

It was found, however, that text availability had little impact on performance, and, in fact, was deleterious in some cases. This leads to the conclusion that memory is not responsible for differences in performance on items testing explicit and implicit comprehension, nor for those found to exist among the four types of inferences. These findings were expected and are consistent with those of Oakhill (1984) who also investigated children's response to explicit and implicit questions in a text-absent and text-present condition. She also concluded that memory could not be held accountable for inferencing difficulties at this age level. The current work not only substantiates Oakhill's findings, but also extends them from the narrative genre to expository text.

Having found that some inferences were more difficult than others and that memory could not explain this difficulty, consideration was given to the third question posed. Could Lipson's results be associated with the saliency of information tested? Could it be that attribute items, for instance, were found easier because the information they focused on was more important to the reader than the information focused on by another type of question? It was assumed that the information salient to readers would appear in recall protocols. Thus protocols were examined to determine if the premise information for some questions was recalled more often than that for other questions? Indeed, the types of questions found easier to recognize were those for which the premise propositions were most

often recalled. On average, however, failure to call premise propositions did not predict performance on recognition items. A reader who did not recall premise information was just as likely as not to answer an inference-based recognition item, although this was less probable for attribute inferences than it was for the other three. Chance level performance was revealed for goal, event and causal inferences. The conclusion to be drawn is that saliency of information tested, as defined by content embedded in recall protocols, does not explain differential performance on inference types.

To this point, then, three conclusions have been drawn. One, Lipson's findings that implicit information is more difficult to process than explicit information and that of the four types of implicit information tested, event and causal inferences are more difficult than attribute and goal inferences have been replicated. Two, neither memory for premises nor, three, saliency of information tested accounts for differential performance. The question then is, how are these differences to be explained? In searching for an answer, reference was made to Frederiksen's theoretical work on inferencing. Specifically, attention was focused on that part of the taxonomy which specifies the size of semantic unit operated upon in generating each type of inference. Table 2 in Chapter Two reveals that the size of the semantic unit varies for each type of inference examined in this study. Those inferences found to be more easily

recognized are those constructed from the smaller semantic units and those found to be more difficult, from the larger, more inclusive units. Whether size of semantic unit predicts the ease with which inferences are generated and recognized needs to be tested directly. Studies employing latency times, for instance, may give some indication regarding the depth of processing required to construct different types of inferences. In such studies, if the hypothesis that larger semantic units are more difficult to use in making inferences, then latencies on the time taken to generate such inferences should be a function of the size of the unit. The benefit of such investigations would be in their ability to provide explanatory rather than simply descriptive data. They may shed light on why certain inferences are more difficult rather than simply stating that differences in degrees of difficulty exist.

This study also sheds light on another question, namely the point at which inferential processing breaks down for items that are more difficult. To address this question, the inferential process can be divided into two components: identifying premise information and constructing the inference from related ideas in the text. The data bearing on these processes would provide information regarding the relationship between question type and the ease with which premise information could be identified. Premise information for explicit questions was identified more readily than that for implicit questions. This is not surprising. It can be explained on the basis

of perceptual matching -- that is to say, words appearing in the correct alternative are identical to those embedded in the passage and the two are simply matched. What is more interesting is that once again students' performance on attribute and goal inferences was superior to that on event and causal inferences. Premise information for the former was correctly identified more often than for the latter. Given that premise information for event and causal items was more difficult to identify, it is not surprising that students fared less well on recognition items based on these premises. If the premises could not be identified as the information necessary to generate the inference, there is little reason to expect that this information would be processed inferentially.

These findings are important because they call into question Lipson's interpretations. To explain why children experienced greater difficulty with inferences in general, and event and causal inferences in particular, Lipson argued that readers were attending to the surface structure of text at the expense of inferential comprehension. Yet the premise identification task in my study revealed differing degrees of difficulty in identifying surface structure information needed to produce the targeted inferences. This finding would suggest that Lipson's explanation is insufficient. While it may account for differences between explicit and implicit information, it leaves unexplained the differences among inference types. Why, for instance, is the surface structure information needed

to construct attribute and goal inferences more readily recognized than that for event and causal inferences?

Lipson's explanation that failure to recognize inference items was a result of readers attending to the surface structure of text was drawn from Tulving and Thomson's (1973) notion of "encoding specificity". She explains that

> what takes place at encoding is what predicts performance at testing. If the cues at retrieval do not match the conception that was formed at encoding, then retrieval will fail. Therefore, if a reader does not generate inferences while reading, inference statements will not be effective retrieval cues at posttesting. (Lipson, 1982, p. 256)

Lipson argued that those items that were more difficult did not serve as viable retrieval cues. As mentioned in Chapter Two, this reasoning suggests that inferences are generated only at the time of encoding. This notion is not supported by my study. In recall protocols, the inferences used as recognition items were rarely included. This might be explained by the notion of production deficiency such that inferences were constructed but simply not reported. If this were the case, however, it would seem reasonable to expect the targeted inferences to be embedded in protocols more often than they were. On the basis of this information it was hypothesized that inferences must be constructed at the time of question answering, an hypothesis made plausible by previous research (Freedle & Hale, 1979; Kintsch, 1974).

The inadequacy of encoding specificity as a basis for

interpreting the present data has thus been demonstrated on two counts. One, it does not explain why there would be discrepant performance among inferences. Two, it fails to consider the possibility that inferences can be constructed at times other than encoding.

The premise identification task not only called into question the explanatory power of encoding specificity. It also focused attention on the need for caution when using a levels of processing argument to predict performance on different types of inferences. A levels of processing argument suggests that because event and causal inferences are "deeper" and therefore more difficult to generate, they would demand greater cognitive processing effort. In turn, this extra processing would increase the likelihood of their being retained in memory, thereby facilitating recognition. However, this line of reasoning was not supported by the results. To explain this situation, it was reasoned that for the more difficult inferences to be generated, it would be necessary to first process the premise information upon which the inferences are based. The data reported above would suggest that this was not happening because seldom was the information needed to construct the inference mentioned in recall. Thus, a levels of processing argument is insufficient to explain the data. The predictive value of this line of reasoning will have merit only when the textual content triggers extra processing required to generate the inference. In this study, the premise information did

not satisfy this criteria. The question is, why not?

For levels of processing to explain performance, overall comprehension of the passage must depend upon generating the theoretically more difficult inference. If a reader's comprehension is not jeopardized by failures to construct such inferences, it would seem unreasonable to assume that extra cognitive effort required would necessarily be applied. What is being referred to here is the centrality of the information tested. The question, simply stated, is how important was the information interrogated? For instance, does understanding of other information in the passage rely on generating the targeted inference? In studies using narrative material it has been found that only those inferences relevant to the progress of the story are made -- relevant inferences being those which are necessary to determine what happened and why (Crothers, 1979; Freedle & Hale, 1979; Warren, Nicholas, & Trabasso, 1979). In reviewing the event and causal items constructed by Lipson, it appears that general understanding of the passage would not be impaired by failure to generate event and causal inferences as the gist of the passage did not depend on their construction. A more stringent way to examine the possible defects of centrality would be to analyze the hierarchical structure of the text using a system such as that developed by Meyer (1975). There are those who believe that factors such as centrality may turn out to be more important determiners of performance than traditional predictors such as high or low order questions (Johnston,

1983). Centrality of information was not controlled in this study and it may have had an effect on children's performance. Research is needed to clarify the effect of lower order questions directed at important information and higher order questions directed at unimportant information (Wixson, 1984).

The Relationship Between Different Measures of Comprehension

There is another purpose to which the data from this study can be applied. It is to analyze the relationship between different measures of comprehension. This study used three methods of assessing comprehension: recall of information following the reading of text; the ability to answer questions about information explicitly presented in text; and the ability to answer questions that call for inferences. Overall, the data suggest that comprehension varies with the method used to assess it. Passage recall, for instance, is a typical measure of comprehension. In this study, it did not correlate with recognition items testing explicit and implicit comprehension. One reason for this may be that information tested by the recognition items was not considered particularly salient by the readers and therefore they did not include it in their recall protocols. The data support this speculation. Moreover, readers did not have to recall premise information in order to answer a recognition item correctly. These findings suggest that measuring comprehension by different methods yields different findings about how well readers comprehend.

This fact must be taken into consideration when comparing results across studies.

Related to the issue of assessing comprehension and the need for caution when comparing results across studies is the issue of the interactions between passages and types of inferences. Though Lipson considered all recognition items of a particular type to be assessing the same ability, these same questions were often the source of significant variance across passages. The reasons for these differences would require empirical investigation, however three possibilities are presented. One, Lipson commented that the location of target inferences had been controlled. However, the results of the premise identification task would suggest that this criterion had not been satisfied, because in many instances, premise information for the same type of question was drawn from different locations in text. It is important to realize that "indicating the part of a text from which a question was derived is <u>not</u> the same as specifying which parts of the text are pertinent to answering it" (Lucas & McConkie, 1980, p. 134). Whether the locational discrepancies of premise information account for interactions between passages and question type would need to be tested directly.

Two, as mentioned in Chapter Two, some questions tested information about which children may have had erroneous preconceived notions. In a portion of Lipson's study not replicated here, she discovered that children often cling to faulty prior knowledge even

when textual content contradicts its validity. This could produce interactions of passage by type of inference as a result of different passages tapping differing degrees of prior knowledge.

Three, an analysis of the incorrect options in each recognition pair raises the question of whether different cognitive operations are required for readers to determine their inappropriateness. For example, some options were directly contradicted by information in the passage. This point is illustrated by the explicit attribute item designed for the Eskimo passage. The text explicitly states that sled dogs are strong and smart. This directly contradicts the information contained in the incorrect alternative which purports that sled dogs are strong, but not very smart. In other cases, the incorrect option appears simply less plausible. The explicit event item constructed for the New Guinea passage demonstrates this point. The text informs the reader that the people of New Guinea still live as people did a thousand years ago. This information would render implausible the incorrect alternative which suggest that tractors are used to grow corn. Whether these factors have affected the current research is unknown. They are mentioned here for two reasons: one, as direction for further research; and two, to demonstrate possible confounds that may affect the validity of the results reported here.

Implications for the Classroom

The above discussion yields information that has important instructional implications. First, it is important to realize that not all questions designed to test the comprehension of implicit information are similar. As Lipson has demonstrated, certain types of inferences are more difficult to process than others. Items designed to test comprehension of an unstated goal are answered more readily than items requiring the construction of an unstated causal relationship. This point should be remembered when designing items for evaluative purposes. If heterogeneity of inference types is not recognized, student strengths and weaknesses may be misdiagnosed. Second, even when the heterogeneity of inference types is taken into consideration, there is still need for caution in developing test items. Support for this comment is garnered from the discussion regarding the different types of incorrect options found in recognition items. If questions are not carefully constructed the cognitive demands of one may differ from those of another, even within inference types, thus jeopardizing accurate diagnosis. The third point relates to the finding that grade three children had greater difficulty identifying premise information for some inferences than others. This leads to the suggestion that in remediating poor performance, instructors should first determine at what point in the process the reader encounters difficulty. Since comprehension involves a complex of mental processes, if comprehension is impaired,

any of a number of factors could be responsible. In planning instruction then it is important to proceed by first identifying which of these factors constitutes the weak link and then developing teaching strategies designed to remediate that weakness. Consider the difficulty grade three children experienced in identifying premise information for event and causal inferences. In this case it would be worthwhile to sensitize students to the sources of information available in text (Raphael, 1982).

From an instructional point of view, it is interesting to note that less skilled readers, who respond to inferential tasks least well, are those who receive the least opportunity to practice these skills (Durkin, 1978-79). Explicit questions are asked of them more frequently than are questions that assess implicit comprehension. As well, material written to satisfy readability criteria for these students also fails to provide an opportunity for practice in generating inferences. In order to simplify text, relational information and connectives are often avoided -- simple sentences are used instead of compound or complex sentences (Marshall, 1981). Unfortunately "if students find the comprehension of groups of words difficult when the connective is absent, they are faced with the greatest percentage of them at the lowest grade level" (Robertson, 1968, p. 397). Children are best at answering the type of question most often asked of them (Hansen, 1981). Thus it is not surprising that less skilled readers remain poorer at answering inference

questions.

Given the importance of inferences to comprehension and the growing awareness of how little time is actually devoted to teaching comprehension of implicitly stated material (Duffy & McIntyre, in press), greater emphasis must be placed on teaching reading comprehension skills (Pearson, 1984). This is especially important for the unskilled reader who is less inclined to generate relationships between groups of words.

This study has revealed that

> the less skilled comprehender's problem does appear to be one of degree, rather than an absolute inability: they <u>can</u> make inferences so it is reasonable to suppose that they would be able to make greater use of such processing if they were more aware of its value. (Oakhill, 1984, p. 36)

and if they were to receive instruction directed to the specific skills found wanting. In assessing those skills it is important to know exactly what it is we are asking the reader to do. As has been illustrated here, inferencing cannot be characterized as a unitary skill. This point must be remembered and reflected in the design of instruction, assessment and remediation.

This point must also be remembered when designing research studies. Anderson (1972) has criticized reading research on the basis that investigators too seldom provide a sufficiently clear description of the experimental task to allow inter-study comparisons. As was

made clear in Chapter Two there are a number of ways of characterizing inferences. Flood (1981) suggests that for the purposes of reading research it may not be necessary to have a standard definition of inferences. However, both he and Anderson, do insist that it is necessary for any reading comprehension study investigating inferences to include a clear description of the type of inference under investigation. Future research on inferential processing should reflect this recommendation.

APPENDIX A

Passages and Recognition Items

This appendix contains the passages and recognition items used in this study. They were developed by Marjorie Lipson, who kindly permitted their use in the current research. Recognition items are listed in the same order, namely, explicit attribute, explicit event, implicit attribute, implicit goal, implicit event, and implicit causal. In each instance, the correct alternative appears first.

Eskimos

People called Eskimos live in the Far North. Winters in the Far North are long and very cold. It is difficult to stay alive. People could die because they have no food or because of the cold. Eskimos have to live together and share whatever they have with their friends.

Many Eskimos hunt for their food. Sometimes they must travel a long way. When the Eskimos stop, they build a house of snow. This snow house is called an igloo. They live in it for awhile. Then they move on to hunt for more food and build new igloos when they stop.

Eskimos travel on sleds pulled by dogs. These dogs are strong and smart. Sleds and dogs are better to use than cars and bicycles in the Far North. When the summer comes, most Eskimo families live in tents. They spend time outdoors getting the meat and furs they will need in the winter. Today many things are changing for the Eskimos. Many live in towns and cities. They buy their food in stores.

Recognition Items: Eskimos

1. Sled dogs are strong and smart.

   Sled dogs are strong, but not very smart.

2. Eskimos lives in tents in the summer.

   Eskimos live in tents when they go hunting.

3. In the summer it is warm in the Far North.

   It is always cold in the Far North.

4. Eskimos often travel long distances to find food.

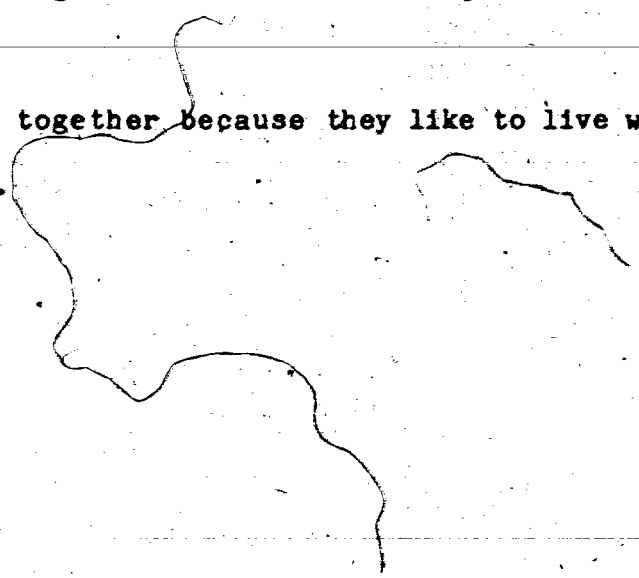
   Eskimos often travel long distances to see friends.

5. Eskimos build many igloos each year.

   Eskimos build only one igloo each year.

6. Eskimos live together because it is dangerous to live alone.

   Eskimos live together because they like to live with their friends.



Venus Flytrap

The Venus Flytrap is a big plant. This plant is different. It can eat animals. The plants are called Flytraps but they catch many kinds of small insects.

Plants do not have tongues, but most of them do have leaves. The leaves on the Venus Flytrap have very small hairs on them. When the hairs are touched, the leaf folds in half. The Venus Flytrap has sweet-smelling sticky leaves. Insects like them. When an insect lands on one of the leaves, it moves around on the leaf and touches the hairs. The leaf closes very quickly. Long points on each side of the leaf look together. The fly is trapped inside. Then the Flytrap eats it. The Flytrap opens up again ready to catch another insect. When you give hamburger to the Venus Flytrap, it will do the same thing.

These plants live in places where the soil is not good. It does not have the things that the plants need to grow. Insects and other animals give the Flytrap the things it needs.

Recognition Items: Venus Flytrap

1. The Venus Flytrap is a plant.

   The Venus Flytrap is an animal.

2. The Venus Flytrap catches flies with its leaves.

   The Venus Flytrap catches flies with its tongue.

3. The Venus Flytrap is a meat-eater.

   The Venus Flytrap is a plant-eater.

4. Flies land on the Venus Flytrap to eat the sticky leaves.

   Flies land on the Venus Flytrap to lay eggs.

5. The Venus Flytrap eats hamburger.

   The Venus Flytrap eats only insects.

6. The Venus Flytrap catches flies because the soil doesn't give them the things they need.

   The Venus Flytrap catches flies because flies would eat it.

The Vikings

Many hundreds of years ago the people of Norway lived in small groups. The vikings were one of these groups. They were strong, brave people who lived near the water. The land was rocky. There was not much good farmland there. The Vikings went to the sea to earn a living.

The men became good ship builders. They built long ships of wood. Some of the ships had sails, but they had men to row them too. Often the ships had carved heads on them.

The Vikings became good fishermen. But later, they were not happy to be just fishermen. They wanted riches. These Vikings became pirates. They attacked ships and robbed many people.

Soon Vikings became very brave about sailing on the sea. They sailed farther than anyone else. They sailed their ships to new places and wrote down what they had seen. One group of Vikings came to the land of America many, many years before Columbus went there. The Indians were not friendly, so the Vikings went back to Norway.

Recognition Items: Vikings

1. The Vikings were fishermen.

   The Vikings were farmers.

2. The Vikings built ships.

   The Vikings built farms.

3. The Vikings were cruel people.

   The Vikings were nice people.

4. The Vikings attacked ships to get riches.

   The Vikings attacked ships to protect themselves.

5. The Vikings discovered America.

   Columbus discovered America.

6. The Vikings went to sea because they couldn't make a living at farming.

   The Vikings went to sea because they wanted to be fishermen.

The People of New Guinea

New Guinea is a large island in the Pacific Ocean. New Guinea people live in tribes. Each tribe is very different but all of them decorate their bodies. It is important to the people to do this. In some tribes there are people called "mud men". They cover themselves with river mud and put on masks made of mud. In other tribes there are "wig men" who wear big headdresses made of human hair.

Many people in New Guinea still live as people did a thousand years ago. They still use sticks to plant their corn and other crops. They grow most of their own food and make the things they need.

Once a year tribes from all over the island come together for a big fair. This fair is called a sing-sing. People bring their best animals and farm crops to show to judges. The judges decide which are the winners.

Then there is dancing and singing. The dancers and singers paint their faces. They wear fur on their ears and headdresses with bright features.

Recognition Items: New Guinea

1. New Guinea wigmen wear wigs made of human hair.

   New Guinea wigmen wear wigs made of feathers.

2. People of New Guinea plant corn with sticks.

   People of New Guinea plant corn with tractors.

3. Sing-sings are happy occasions.

   Sing-sings are sad occasions.

4. Tribesmen bring animals to fair to win prizes.

   Tribesmen bring animals to fair to sell them.

5. New Guinea people make a living by farming the land.

   New Guinea people make a living by singing and dancing.

6. Tribesmen dress up because they like to decorate themselves.

   Tribesmen dress up because they want judges to decide the winners.

APPENDIX B

Contained in this appendix is a description of the preliminary study conducted to determine the passage dependency of the recognition items developed by Lipson.

A preliminary study was conducted to determine the passage dependency of the recognition items developed by Lipson. It was conducted approximately six weeks prior to the experimental session and involved the same children who were to participate in the study. The experimenter read aloud each pair of recognition items for each of Lipson's eight passages: Eskimos, Venus Flytrap, The Vikings, Dinosaurs, Insects, Whales, The People of New Guinea, and Totem Poles. Students did not read the passages prior to answering the questions. Participants were instructed to listen and to mark which of the two sentences they considered correct. Should they be unable to decide which of the sentences was correct, they were advised to indicate their uncertainty by placing a question mark in the appropriate space. It was emphasized that this response was preferable to guessing. For each passage, presentation order of the items was determined by tossing a die. The arrangement of paired statements was decided by flipping a coin.

Confidence intervals were calculated and revealed an upper limit of 21.5. Therefore, in the pre-testing condition, any question which received 22 or more correct responses was considered inappropriate as this indicated the ability to respond correctly did not depend upon reading the passage.

In determining which passages to include in the study, it was decided to focus on the number of correct responses to causal items. This was done for the following reasons. There was considerable

response variability across passages and question type and the interactions would be difficult to unravel. Therefore, the decision was made to focus on one type of question. Causal inferences were chosen. They were of particular interest because Lipson's research had identified them as more difficult.

For three of Lipson's eight passages, the children's ability to answer causal questions correctly reliably exceeded chance. For Dinosaurs, Totem Poles and Insects, the correct alternative was selected by 24, 23, and 22 participants respectively. On the basis of these findings, these passages were eliminated.

Time constraints necessitated deleting a fourth passage. The preliminary study had shown that of the five remaining passages considered acceptable, Whales contained content with which the children were most familiar. Thus, it too was eliminated.

APPENDIX C

This appendix lists examples of statements embedded in recall protocols which were considered semantically equivalent to those appearing in the stimulus passages. In each case, the recalled information follows that which appeared in the text.

Eskimos

1. Today many things are changing.

   But now these days they would ...

   But most of the Eskimos now ...

   Nowadays...

2. They must share...

   They have to share...

3. They build new igloos.

   Again they make igloos.

4. Then they move on and build new igloos.

   The build new igloos wherever they go.

New Guinea

1. The singers and dancers paint their faces.

   The singers and dancers dress their faces up.

2. New Guinea people live in tribes.

   In New Guinea there's certain tribes.

3. Tribes ... come together for a fair.

   They have a festival.

4. People bring their best animals and farm crops to show to judges.

   They have something like a contest.

5. Then there is singing and dancing.

   Then they started to have a party.

6. They cover themselves with mud.

   They dress up with mud.

7. Judges decide the winner.

   The judge sees who wins.

8. They grow most of their own food.

   They do crops. They have many crops.

Venus Flytrap

1. The leaf closes.

   The leaf rolls up.

2. The fly is trapped.

   And the bug that's inside there is helpless.

3. These plants live where the soil is not good.

   It lives - where the soil - barely any other plants grow.

4. Insects and other animals give the Flytrap the things it needs.

   They have to eat bugs that come on it or they will die.

Vikings

1. They sailed their ships to new places.

   They went to places that other people had never been.

2. They became shipbuilders.

   They built boats.

3. The Indians were not friendly.

   The Indians didn't like it.

4. They became fishermen.

   They went fishing.

REFERENCES

Anderson, R. C. (1972). How to construct achievement tests to assess comprehension. Review of Educational Research, 42, 145-170.

Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. Cambridge, England: Cambridge University Press.

Black, J. G. (1983, December). An exposition on understanding expository text. Cognitive Science Technical Report No. 28. Yale University, Cognitive Science Program.

Bormuth, J. R., Carr, J., Manning, J., & Pearon, P. D. (1970). Children's comprehension of between and within sentence syntactic structures. Journal of Educational Psychology, 4, 349-357.

Bovair, S., & Kieras, D. E. (1981, July). A guide to propositional analysis for research on technical prose. (Tech. Rep. No. 8). (Contract No. N00014-78-C-0509) Office of Naval Research.

Bridge, C., Tierney, R., & Cera, M. (1977, December). Inferential operations of children involved in discourse processing. Paper presented at the National Reading Conference, New Orleans.

Broatch, C. M. (1981). Inferencing and reading comprehension. Unpublished manuscript, Simon Fraser University, Faculty of Education.

Clark, H. H. (1977). Inferences in comprehension. In D. LaBerge & S. J. Samuels (Eds.), Basic processes in reading: Perception and comprehension (pp. 243-264). Hillsdale, N.J.: Erlbaum.

Crothers, E. J. (1979). Paragraph structure inference. Norwood, N. J.: Ablex.

Drum, P. A., Calfee, R. C., & Cook, L. K. (1981). The effects of surface structure variables on performance in reading comprehension tests. Reading Research Quarterly, 16, 486-514.

Duffy, G., & McIntyre, L. A. (in press). A naturalistic study of instructional assistance in primary grade reading. Elementary School Journal.

Durkin, D. (1978-79). What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 14, 481-533.

Englert, C. S., & Hiebert, E. H. (1984). Children's developing awareness of text structure in expository materials. Journal of Educational Psychology, 76, 65-74.

Flesch, R. A. (1955). Why Johnny can't read. New York: Harper & Row.

Flood, J. A. (1978). A system for scoring inference from recall data. National Reading Conference Yearbook. Clemson, S.C.: National Reading Conference.

Flood, J. A. (1981). Prose comprehension: A selected review of literature on inference-generation as a requisite for understanding text. In D. F. Fisher & C. W. Peters (Eds.), Comprehension and the competent reader: Inter-specialty perspectives (pp. 51-68). New York: Praeger Publishers.

Frederiksen, C. H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 7, 371-458.

Frederiksen, C. H. (1977). Semantic processing units in understanding text. In R. O. Freedle (Ed.), Discourse in production and comprehension (pp. 57-87). Norwood, N.J.: Ablex.

Frederiksen, C. H. (1979). Discourse comprehension and early reading. In L. G. Resnick & P. A. Weaver (Eds.), Theory and practice of early reading: Vol. 1 (pp. 155-186). Hillsdale, N.J.: Erlbaum.

Freedle, R. O. (Ed.). (1977). Discourse processes: Advances in research and theory. Vol. 1. Discourse production and comprehension. Norwood, N.J.: Ablex.

Freedle, R. O. (Ed.). (1979). Advances in discourse processes: Vol. 2. New directions in discourse processing. Norwood, N.J.: Ablex.

Freedle, R. O., & Hale, G. (1979). Acquisition of new comprehension schemata for expository prose by transfer of a narrative schema. In R. O. Freedle (Ed.), New directions in discourse processing (pp. 121-135). Norwood, N.J.: Ablex.

Gallagher, M. C., & Pearson, P. D. (1982, December). An examination of expository texts in elementary instructional materials. Paper presented at the annual meeting of the National Reading Conference, Clearwater, Fl.

Hansen, J. (1981). The effects of inference training and practice on young children's comprehension. Reading Research Quarterly, 16, 391-417.

Johnson-Laird, P. N. (1983). Mental models: Towards a cognitive science of language, inference, and consciousness. Cambridge: Cambridge University Press.

Johnston, P. H. (1983). Reading comprehension assessment: A cognitive basis. Newark, Delaware: International Reading Association.

Kintsch, W. (1974). The representation of the meaning in memory. New York: John Wiley & Sons.

Lehnert, W. G. (1981). A computational theory of human question answering. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), Elements of discourse understanding (pp. 145-176). Cambridge: Cambridge University Press.

Lipson, M. Y. (1982). Learning new information from text: The role of prior knowledge and reading ability. Journal of Reading Behavior, 14, 243-261.

Lucas, P. A., & McCopnkie, G. W. (1980). The definition of test items: A descriptive approach. American Educational Research Journal, 17, 133-140.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. Child Development, 48, 986-992.

Marshall, N. (1981). The application of basic research to reading instruction. In D. F. Fisher & C. W. Peters (Eds.), Comprehension and the competent reader: Inter-speciality perspectives. New York: Praeger.

Marshall, N. (1977, December). Inferred and generated information in freely recalled text. Paper presented at the National Reading Conference, new Orleans.

Meyer, B. J. F. (1975). The organization of prose and its effect on memory. Amsterdam: North-Holland Publishing Co.

Meyer, B. J. F., & Rice, G. E. (1984). The structure of text. In P. E. Pearson (Ed.), Handbook of reading research (pp. 319-351). New York: Longman, Inc.

Oakhill, J. (1982). Constructive processes in skilled and less-skilled comprehenders' memory for sentences. British Journal of Psychology, 73, 13-20.

Olson, M. W. (1985). Text type and reader ability: The effects on paraphrase and text-based inference questions. Journal of Reading Behavior, 17, 199-214.

Paris, S. G. (1975). Integration and inference in children's comprehension and memory. In F. Restle, R. M. Shiffrin, N. J. Castellan, H. R. Lindman, D. B. Pisoni (Eds.), Cognitive theory (Vol. 1). Hillsdale, N.J.: Erlbaum.

Paris, S. G., & Lindauer, B. K. (1976). The role of inference in children's comprehension and memory for sentences. Cognitive Psychology, 8, 217-227.

Paris, S. G., & Upton, L. R. (1976). Children's memory for inferential relationships in prose. Child Development, 47, 660-668.

Pearson, P. D. (1984). Direct explicit teaching of reading comprehension. In G. G. Duffy, L. R. Roehler, & J. Mason (Eds.), Comprehension instruction: Perspectives and suggestions (pp. 222-233). New York: Longman.

Pearson, P. D., & Johnson, D. D. (1978). Teaching reading comprehension. New York: Holt, Rinehart, and Winston.

Raphael, T. (1982, March). Improving question-answering performance through instruction. (Reading Education Report No. 32). Campaign, Illinois: University of Illinois at Urbana-Champaign, Center for the Study of Reading.

Reder, L. M. (1980). The role of elaboration in the comprehension and retention of prose: A critical review. Review of Educational Research, 50, 5-53.

Robertson, J. E. (1968). Pupil understanding of connectives in reading. Reading Research Quarterly, 111, 387-417.

Schank, R. C. (1975, June). Using knowledge to understand. In R. C. Schank & B. L. Nash-Webber (Eds.), Theoretical issues in natural language processing. Proceedings of the International Joint Conference on Artificial Intelligence, MIT, Cambridge, Mass.

Spiro, R. J. (1980). Constructive processes in prose comprehension and recall. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension (pp. 245-278). Hillsdale, N.J.: Erlbaum.

Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. Journal of Educational Research, 8, 323-332.

Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Reader, 80, 352-373.

Waller, T. G. (1976). Children's recognition memory for written sentences: A comparison of good and poor readers. Child Development, 47, 90-95.

Warren, W. H., Nicholas, D. W., & Trabasso, T. (1979). Event chains and inferences in understanding narratives. In R. O. Freedle (Ed.), New directions in discourse processing (Vol. II) (pp. 23-52). Norwood, N.J.: Ablex.

Wilson, M. M. (1979). The processing strategies of average and below average readers answering factual and inferential questions on three equivalent passages. Journal of Reading Behavior, 11, 235-245.

Wittrock, M. C. (1984). Generative reading comprehension. Unpublished manuscript. University of California, Los Angeles, Graduate School of Education.

Wixson, K. K. (1984). Level of importance of postquestions and children's learning from text. American Educational Research Journal, 21, 419-433.